The Neighborhood Effect on Health Outcomes for Women in Urban India

Priyanka Venkannagari, Duke University

Dr. Charles Becker, Faculty Advisor
Dr. Kent Kimbrough, Seminar Advisor
Dr. Michelle Connolly, Seminar Advisor

Honors thesis submitted in partial fulfillment of the requirements for Graduation with Distinction in Economics in the Trinity College at Duke University.

Duke University
Durham, NC 27708
2017

Acknowledgements

I would like to thank Professor Charlie Becker for his patience and support in the development of this paper. Thank you for taking the time to edit various versions of this paper, for providing constructive criticism, and for making the process more manageable. I would also like to thank Professor Kimbrough and Professor Connolly for their comments, feedback, and time. I would also like to thank everyone in my thesis seminars for their input, ideas and help with the process.

Abstract

The paper uses 2011 Indian Human Development Survey data to assess the impact of 5 categories of variables on health outcomes. It uses OLS models, interaction terms, instrumental variable models, fixed effects and random effects to investigate the existence of a neighborhood effect on health outcomes for women in urban India. This paper finds that various aspects of health practices, empowerment, amenities and financial security are relevant when looking at health outcomes. Interventions looking to address health outcomes should consider these variables and the compounding neighborhood effect.

JEL Classification: C36; I1; I12; O18

Keywords: Health; India; Women; Urban

Table of Contents

I. Introduction

The neighborhood effect on health, education, and crime has been studied extensively in higher-income countries. There is a broad set of research that has looked at the way neighborhood environments can create impacts beyond individual level determinants. This study provides a comprehensive look at the neighborhood level impacts on health outcomes for women in urban India. I want to understand how much of the variation in health outcomes is accounted for at the neighborhood level versus at the individual level. This paper will introduce the broader context of the health environment in India, provide an in-depth look at the current literature on women's health in India and neighborhood effects on health in low-income countries, and then use an extrapolated methodology to test for the existence of neighborhood effects.

We can see the implications of neighborhood level health through studies conducted in higher-income nations. With the rise of obesity in the United States, researchers have begun studying the relationship between built environments, physical activity and other health outcomes. Obesity has been linked to the physical attributes of neighborhoods in the US. Both lower and higher income groups benefit from living in walkable neighborhoods, when controlling for income (Sallis *et. al.* 2009). Research finds that the relationship between neighborhoods and health outcomes persists throughout the life cycle of individuals (Menec *et. al.* 2010). Lower-income countries are likely experiencing similar neighborhood effects, albeit in a different context, but the research is less extensive on the subject.

India has urbanized rapidly in the past decade, with more and more people moving to densely populated neighborhoods in booming urban spaces. This is simultaneously a sign of growth and a reason for concern. India's rapid urbanization has coincided with a shift from infectious diseases towards more chronic illnesses. India is currently in the middle stage of this

transition with growing incidences of chronic illness and the persistence of infectious diseases (Barik and Desi 2013).

In comparison to Brazil, Russia, and China, India has the lowest per capita spending on health and simultaneously has the worst health outcomes (Barik and Desi 2013). India exhibits high out of pocket spending on health care, with a large proportion of households facing health payments that exceed 10% of household expenditures in 2004-2005 (Barik and Desi 2013, Ghosh 2011). Regional variability in health spending is high, with 3.46% of households in Assam spending more than 10% of their income on healthcare and 32.42% of households in Kerala spending more than 10% on healthcare (Ghosh 2011, Barik and Desi 2013).

The government has made progress towards improving health infrastructure through their 11th Five-Year Plan and the 12th Five-Year Plan. These plans promise significant improvements in India's health system. The promise of universal health coverage was made in the 12th Five-Year Plan (Barik and Desi 2013). While the actual outcomes of this plan are yet to be determined, this marks a huge step for the Indian health system. Currently, about 2/3 of the population seeks private care, due to lack of trust in government hospitals. The utilization of public health services is still extremely low in all socioeconomic groups. There is a shortage of health professionals and a persisting assumption that public health is of lower quality (Barik and Desi 2013). Addressing health issues is an important part of the puzzle as India continues to develop.

Improving health outcomes in India will take time and investment. There is a significant need to understand how neighborhood effects can impact the health outcomes interventions are targeting. This study aims to fill the gap in the literature and focus specifically on the neighborhood effect on female health outcomes in urban India.

II. Literature Review

There are three relevant strands of literature pertaining to my topic: variables impacting health outcomes, regional differences in health in India, and the neighborhood effect on health.

i. Variables impacting health outcomes

Topic	Author(s)	Main Findings
Financial Assets	Beck, Pulkki-Brannstorm, San Sabestian 2015 Clayton, Liiares-Zegarra, Wilson 2015 Sood *et al.* 2014 Sweet, Nandi, Adam, Mcdade 2013	Studies find that base income significantly reduces odds of minor illness, while debt increases incidences of depression, worse self reported heath and higher blood pressure. Specifically in India, Sood *et al* 2014 finds that insuring poor households for costly and underused health services significantly improves population health.
Health Practices	Aiello *et al* 2008 Chhaploa, Brar 2015 Dobe, Mandal, Jha 2013 Mainassara, Tohon 2014 Nguyen, 2012 Spears 2013	Poor water quality and open defecation can have detrimental impacts on health outcomes such as childhood stunting. Hand washing is associated with maternal good hand washing practices, presence of sanitary latrine, availability of soap at hand washing locations and income. Hand washing has been found to control infections. Improving health infrastructure directly impacts hygiene related habits.
Household Amenities	Halpenny *et al* 2012 Guo, Zhang, Sherraden 2009 Nguyen, 2012	Literature finds that household assets and surroundings impact health. In China, Guo 2009 finds assets directly affect access to medical care and indirectly affect health by influencing health behavior and psychological condition. Other studies also show similar results, that with increasing assets health outcomes also improve.
Female Empowerment	Garces-Ozanne, Kalu, Audas 2016 Laverack 2006 Santhya *et. al* 2010	Studies find that higher levels of empowerment have positive associations with life expectancy for females and interventions in developing countries frequently use empowerment as a mechanism to improve health. Marriage age also has implication on health; marriages that occur at older ages are associated with higher rates of rejecting wife beating, using contraceptives, and having their first birth in a health facility

ii. Regional differences in health outcomes and healthcare utilization in India

District level differences in health care in India exist.[1] Desi and Lujian showed that 47% of the variation in delivery of care is between districts, while 53% is within districts, showing that most variation exists within a singular district. The urban-rural divide on health outcomes is also great. However, utilization of health care is poor in both urban and rural areas (Vargase 2013).

Numerous studies have also looked at the health outcomes in slum vs. non-slum areas. Ghosh (2014) finds that the negative impact of air pollution on health is three times the magnitude for slum children than for children growing up in non-slum areas. Ghosh (2012) finds that children in slum areas are more susceptible to acute and chronic health risks in comparison to non-slum children. The types of slums also affect health outcomes. Subbaraman (2012) finds that outcomes for those living in slums not recognized by the government are worse than those in living in slums recognized by the government. Kulkami (2012) compares two neighborhoods of similar socio-economic statuses (SES) that have different living conditions. The study finds that individuals living in less densely packed housing with better access to amenities are at lower risk of chronic illnesses.

Slum updating has proven to be successful, at least in reducing water-borne illnesses (Butala *et. al.* 2010). Wankhade (2015) finds that open defecation is minimal in cities with one million plus population, but is a much bigger issue in smaller cities. The author uses the availability of public funds to explain this difference. Larger cities get larger amounts of public funding and much of this funding has been invested in sanitation projects (Wankhade 2015).

The literature on regional differences in health outcomes is robust, but few empirical studies exist on district and neighborhood level difference. Many studies on regional differences are

[1] Districts are administrative division within a State and some districts are further divided into

focused on specific aspects of health like sanitation or air-borne illnesses; rarely do they measure overall health outcomes.

iii. Neighborhood effects

Arcaya *et al.* (2016) provide a thorough literature review on all of the neighborhood effects research done over the past 20 years in the United States. They analyze 256 multilevel neighborhood and health papers published between 1995 and 2014. 71% of the studies are cross-sectional and only 20% are longitudinal. Socioeconomic status, built environment, and poverty are the most often-studied neighborhood characteristics. The most often-studied health outcomes are obesity/BMI and mental health. 96% did not acknowledge limitations of their spatial models in regard to how areas are defined. The researchers state that moving forward in the US context, studies should focus on casual models, improving the definition of neighborhoods, and identifying interventions that work (Arcaya *et al.* 2016). The neighborhood effect on outcomes, health and non-health, has been extensively studied in the US setting.

In contrast, the literature on neighborhood effects in developing country is just beginning to improve. Only two major multilevel studies exist: Ahmad *et. al.* 2013 and Ackerson et. al 2008. Ahmad *et. al.* (2013) focuses on the impact of neighborhoods on self-related health in Aleppo, Syria. This study uses a multilevel regression model to investigate neighborhood associations and finds that women living in informal neighborhoods are less likely to report worse self-reported health than women in formal neighborhoods. They find that the neighborhood average of household item ownership is associated with an increase in excellent self-reported health in women.

Ackerson *et. al* (2008) come the closest to the question addressed in this paper. The authors study variation of underweight and overweight women in India at the neighborhood level

by using a multilevel model to calculate neighborhood and state variation of nutritional status. They find substantial variations in overweight and underweight women at the neighborhood and state level. These effects are not captured when looked solely at the individual levels. This study shows the need for interventions to address neighborhood characteristics in order to be successful. This study provides motivation for this study. Instead of looking at neighborhood effects on weight, this paper focuses on more generic measures of health. Ackerson *et. al.* (2008) note that there is a general lack of research on neighborhood influences on nutritional status in low-income countries, which provides the major motivation for my study: to help fill the gap of research knowledge that exists in this space.

III. Empirical Specification

This paper studies the impact of indicator variables in five main categories (demographic, health, empowerment, household assets, and financial security) on health outcomes. This analysis explores these relationships in five parts: linear OLS, OLS with interaction terms, instrumental variables, fixed effects, and random effects regressions.

OLS regression allows us to estimate the impact of the indicator variables on health outcomes. These regressions allow one to look at these relationships when not considering neighborhood level variations. The results of these models will expose issues of endogeneity. The interaction terms and instrumental variable models are constructed to tackle the endogeneity issue and construction of the variable is discussed in the results section.

Multi-level studies are widely used to study neighborhood effects on health outcomes both in the US and in lower income countries. This study uses multilevel regressions, similar to Ackerson *et al.* (2008), to study the neighborhood effect of health outcomes. Multi-level

regressions allow you to see if variability can be attributed to levels of groups such as neighborhoods and cities. The following equation is used in the fitting of this model:

$$y_{ijk}=(\beta_0+\beta X)+v_i+u_{ij}+e_{ijk}$$

y_{ijk} = Health outcomes

$\beta_0+\beta X$ = Indicator variables

v_i = Random effects attributed to cities

u_{ij} = Random effects attributed to neighborhoods

e_{ijk} = Random effects attributed to individuals

If *u* and *v* are substantial, that shows the existence of variation at those levels. Random effects are stronger than fixed effects. Random effects take into account the hierarchy that the populations are sampled from (*i.e.* city, neighborhood). Random effects vary across individuals as well. The majority of studies involving neighborhood effect of health outcome use random effects as the primary mechanism to study these effects. Random effects are preferred because they allow us to study the clustering of health outcomes that might arise due to unmeasured neighborhood and city characteristics. Using random effects also allows us to partition total residual variance into variance between cities and neighborhoods and within neighborhoods.

IV. Data

This analysis uses the Indian Human Development Survey – II (2011). The survey was conducted in 42,152 households across 1503 villages and 971 urban neighborhoods in India. The survey was done in two rounds: 2005-4 and 2011-12. This study used the 2011 instead of the 2005 data set due to the existence of a greater variety of variables. The surveys were conducted

as two one-hour interviews. The data provides neighborhood designations for households in India's 6 largest metro areas: Mumbai, Delhi, Kolkata, Chennai, Bangalore, and Hyderabad. The size of each neighborhood unit differs; which is a potential limitation for this study. This research focuses on women interviewed in these six urban areas. The IHDS is split into three separate surveys with distinct questions: a household survey, an individual survey, and a woman-specific survey. Dependent and explanatory variables are taken from all three surveys.

Table 1. Summary of Cities and Neighborhoods in IHDS

City	**# Neighborhoods**	**N**
Mumbai	24	432
Delhi	39	876
Kolkata	21	780
Chennai	12	190
Bangalore	24	246
Hyderabad	9	320
Total	**129**	**2,884**

i. Dependent Variables

The dependent variables are different indicators of health: self-reported health and days ill. Table 2 shows the summary statistics for the available variables. The self-reported health question asks women to rate their own health on a range from very good (1) to very poor (5). This variable has been widely used in health studies but still comes with its limitations. It is a poor measure of health because of various levels of tolerance and subjectivity associated with self-reporting. Days ill represents the number of days the respondent was ill in the past 30 days. Both variables are heavily skewed to the left, indicating good health on an average day. The means of each of the dependent variables vary by city and by neighborhoods within a city, providing initial motivation that neighborhood effects exist.

Table 2. Summary Statistics of Potential Dependent Variables

Name	Mean	Std. Dev.	Min	Max
Self Reported Health	2.089	0.772	1	5
Days Ill (Past 30)	1.203	3.172	0	30

Table 3 shows correlations between the potential dependent variables. Self-reported health and days ill are not highly correlated. This is possible due to the fact that days of illness reveal short-term illnesses while self-reported health reveals more about long-term health outcomes. Nonetheless, it suggests substantial subjectivity in one or both measures. Self-reported health specifically has been both criticized and praised as a measure of health. Criticisms focus on the fact that self-reported health is a subjective measure of health and could vary based on individual perceptions of health.

Table 3. Correlation Coefficients of Dependent Variables

	Self-Reported Health	Days Ill
Self Reported Health	1.0000	
Days Ill	0.1704	1.0000

ii. Explanatory Variables

This research focuses on explanatory variables in 5 distinct categories: basic demographics, health practices and indicators, female empowerment, household assets, and financial security. These variables were chosen based on their potential impact on health outcomes that are observed in previous literature. Assets and education are the only two variables with a significantly high correlation of about 0.60. Exact values of explanatory variables, constructed variables explanations, and a correlation chart can be found in the Appendix.

Since increasing self-reported health and days ill are associated with worse health outcomes, a predictive positive sign on an indicator variable suggests that increasing a variable is associated with worse wealth outcome. A negative sign suggests that decreasing that variable is associated with better health outcomes.

The demographic variables are self-explanatory. Possible responses for the hand washing and water purification questions are never (1), sometimes (2), usually (3), and always (4). Increasing values are associated with increasing frequency of hand washing and water purification. Community crime is a dummy with one associated with the existence of crime in the area. Own health is also a dummy with one indicating that the respondent has the most say in decisions about their own health. Individual freedoms is an index of six separate questions; increase value is associated with the respondent have more personal freedom.[2] The assets variable is an index constructed by the Indian Human Development survey that includes a combination of 30 household items. The negative surroundings variable is a dummy with one associated with the existence of stagnant water or feces outside of the household. All the other financial variables outside of income are dummies.

[2] Exact questions are listed in the appendix

Table 4. Summary Statistics of Explanatory Variables

Category	Name	Description	Constructed	Mean (St. Dev.)	Range	Predicted Sign
Demographic variables	Age	Individuals ages		37.873 (8.928)	15 – 65	+
	HH size	# people in household		4.873 (1.908)	1 – 22	+
	Education	Education level		7.918 (5.146)	0 – 16	-
	Caste/Religion	Dummy for each cast				
	Marriage Age	Age of marriage		18.535 (3.581)	5 – 39	-
Health Practices	Hand washing	Hand wash after defecation?		3.724 (0.536)	1 – 4	-
	Water Purification	HH purify water before drinking?		1.927 (1.190)	1 – 4	-
Empowerment	Community Crime	Incidence of crime in area	x	0.066 (0.248)	0 – 1	+
	Own Health	Control over own health		0.168 (0.374)	0 – 1	-
	Individual Freedom	Level of individual freedoms	x	4.278 (1.153)	0 – 6	-
Household Assets	Assets	Index of 30 diff assets a household has	x	19.980 (5.458)	1 – 33	-
	Negative Surroundings	Existence of feces or still water near HH	x	0.187 (0.390)	0 – 1	+
Financial Security	Income	HH Income		191676 (228086)	-48460 – 4247000	-
	Debt	Is HH in Debt?		0.462 (0.499)	0 – 1	+
	Life insurance	Does HH have life insurance?	x	0.495 (0.500)	0 – 1	-
	Home owner	Does HH own their home?		0.755 (0.430)	0 – 1	-

V. Results

Five distinct models are used to analyze the data: linear OLS, OLS w/ interaction terms, instrumental variables, fixed effects and random effects regressions. Table 5 shows all the regressions for self-reported health. Model 1 shows the OLS regression, which has an R^2 of 0.076. Hand washing, water purification, individual freedoms, assets, surrounding and home ownership are all significant at the 1% level, although hand washing and home ownership exhibited opposite directionality of what was expected. Hand washing, home ownership and negative surrounding have the largest magnitude. Water purification, individual freedoms, assets, and negative surroundings all exhibit expected signs. Individuals who purify their water more frequently respond with better self-reported health. Increasing household assets and individual freedoms are also associated with better self-reported health, while negative self-reported health is associated with the existence of still water and feces in the individual's neighborhood. In comparison to other significant variables, negative surroundings have the largest magnitudes. This suggests that the environments that respondents live in are essential in understand health.

The only financial variable that is not significant is life insurance, which suggests that the other financial variables might be more powerful indicators of better health outcomes. Health insurance might be a more important measure to look at. Studies should look at the true effects of increasing access to health insurance and the impact it may have on individual's health. These results suggest that different variables within each category (excluding demographic variables) effect self-reported health. Interventions looking at improving health outcomes should consider impacting one of the categories as a mechanism for change.

Table 5. Self-Reported Health Regression Results

Category	Variables	Model (1) OLS	Model (2) Interaction Term	Model (3) Instrumental Variables	Model (4) City Fixed Effects	Model (5) Random Effects
Demographic	Age	0.036 (0.057)	0.037 (0.057)	0.668 (0.432)	0.058 (0.057)	0.050 (0.055)
	Age^2	-0.001 (0.002)	-0.001 (0.001)	-0.025 (0.015)	-0.001 (0.015)	-0.001 (0.001)
	Age^3	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
	HH size	-0.002 (0.008)	-0.002 (0.008)	0.006 (0.012)	-0.002 (0.007)	-0.004 (0.008)
	Education	-0.005 (0.004)	-0.005 (0.004)	-0.013 (0.007)	-0.006 (0.004)	-0.005 (0.004)
	Caste/Religion	Insignif.	Insignif.	Insignif.	Insignif.	Insignif.
	Marriage Age	-0.007 (0.005)	-0.007 (0.005)	0.004 (0.007)	-0.006 (0.005)	-0.004 (0.004)
Health Practices	Hand washing	0.147*** (0.027)	0.031 (0.063)		0.052* (0.029)	0.054* (0.029)
	Hand washing IV			0.555*** (0.147)		
	Water Purification	-0.042*** (0.013)	-0.281** (0.117)	-0.033* (0.019)	-0.015 (0.013)	-0.018 (0.013)
	Interaction		0.062** (0.030)			
Empowerment	Community Crime	-0.031 (0.056)	-0.034 (0.056)	-0.054 (0.094)	-0.003 (0.056)	0.017 (0.055)
	Own Health	0.057 (0.038)	0.055 (0.038)	-0.036 (0.063)	0.079** (0.039)	0.057 (0.038)
	Individual Freedom	-0.046*** (0.013)	-0.047*** (0.012)	-0.023 (0.019)	-0.061*** (0.013)	-0.056*** (0.013)
HH Assets	Assets	-0.011*** (0.004)	-0.011*** (0.004)	-0.018*** (0.006)	-0.002 (0.004)	-0.005 (0.004)
	Negative Surroundings	0.155*** (0.037)	0.159*** (0.037)	0.141*** (0.000)	0.140*** (0.037)	0.173*** (0.037)
Financial Security	Income	0.000** (0.000)	0.000** (0.000)	0.000 (0.000)	0.000*** (0.000)	0.000* 0.000
	Debt	0.055* (0.029)	0.055* (0.029)	0.103** (0.046)	0.037 (0.031)	0.039 (0.030)
	Life insurance	-0.018 (0.031)	-0.016 (0.031)	0.013 (0.048)	-0.051 (0.032)	-0.046 (0.031)
	Home owner	0.118*** (0.033)	0.115*** (0.033)	0.040 (0.607)	0.098*** (0.033)	0.070** (0.034)
City Dummies	Delhi				0.116** (0.057)	
	Kolkata				0.267*** (0.058)	
	Chennai				-0.265*** (0.072)	
	Constant	1.302 (0.852)	1.742 (0.877)	-5.998 (3.977)	1.280 (0.839)	1.318 (0.817)
	Adjusted R^2	0.076	0.077	.	0.114	.
	N	2,844	2,844	1,269	2,844	2,844

*** p<.01, **p<0.5, * p<0.1

Hand washing and home ownership exhibit the opposite directionality of what is expected, namely worse health outcomes. The literature shows that quality of housing, landlord relations, and positive wider neighborhood relations can make renting a house a positive influence on health (Clarke 2012). The causal link is unclear, for example owner occupied may be in areas with poorer infrastructure. The data provides an explanation for this result; 75% of the respondents own their own homes, which is a relatively high percentage of people. Looking at the raw data it seems as if at higher asset levels, there are a greater percentage of people who are renting. This would suggest that wealthier individuals are more likely to rent than poorer individuals.

We would expect that increasing the frequency of hand washing would be associated with better health, but there could be some endogeneity issues. For example with hand washing, it is likely that a person who is sick starts washing their hands more often. Most literature on hand washing focuses on how it controls specific types of infections (Aiello et al 2008, Chhaploa 2015). Since the dependent variables in this paper measure overall generic health, it might be less likely to capture that. There might be an interaction between location of sink in the household and use of soap, which could not be studied due to limitations in the dataset. It could also be because poorer neighborhoods are exposed to more interventions focused in increasing hand washing. For these reasons Models 2 and 3 focus on using interaction terms and instrumental variables to explore the relationships between neighborhood hand washing norms with individual health outcomes.

Model 2 uses an interaction term that interacts hand washing with water purification. It is possible that hand washing alone might not impact health, but also the quality of the water used to wash hands. We would expect that those who wash hands more frequently and purify their

water more frequently would exhibit better health. All variables that are significant in Model 1 are still significant except hand washing. However water purification is still significant and so is the interaction term. The results of the interaction term are still surprising and suggest that self-reported health is worse among those who hand wash more frequently and purify more frequently in comparison to individuals who do neither. Hand washing was also interacted with various other terms including assets but in those models the interaction terms were insignificant.

Instrumental variables (IV) are used when there are potential endogeneity issues. In this case, those who are sick might be washing their hands more frequently due to their illness. Ideal IV's are variables that are strong predictors for the endogenous variables but are uncorrelated with the independent variable and error term. Model 3 uses an instrument that was constructed by splitting the data set into women above the average age of 37 and women below that age. Then the IV was constructed using the average rates of hand washing for women above 37 for each neighborhood. The averages of the older women are used in place for the younger women and the older women are left out of the regressions. The constructed IV is a strong instrument, because we can reject the null hypothesis that the instrument is weak at the 5% significant level using both the two-stage least regression and LIML Wald statistic. A two stage least squared regression was then used to arrive at the results in Model 3. The instrument still reveals that increasing frequency of hand washing is associated with poorer self-reported health. More specifically, individuals living in neighborhoods with higher hand washing frequencies have worse health outcomes.

Both the interaction term model and the instrumental variable model are unable to explain the hand washing results. There could be various explanations of this outcome. There could be reporting errors in the data; 76% of respondents stated that they always wash their

hands and only one respondent stated that she never washed her their hands. Looking at hand washing responses relative to assets levels reveals no significant patterns that could explain these results. We expect hand washing practices would differ based on socio-economic status, but this does not seem to be the case. The high percentage of respondents who reported that they always wash their hands could be reducing the significance of the results. Accuracy of responses is a major issue with survey data, and inaccuracies in reporting could be causing these results. Ideally we would also interact hand washing with other terms. A potential concern with the IV model is that the use of hand washing among older women in the neighborhood may be a good fist stage regression but still ultimately still correlated with the dependent variable. An endogeneity issue might still exist, and testing alternative IV's could reveal a more through understanding of the relationship between hand washing and health.

Model 4 and 5 focus on the regional differences in health outcomes; model 4 uses city fixed effects, while model 5 uses random effects. Fixed effects and random effects are both ways to study the impact of location on health. Due to the reasons previously mentioned, random effects are preferred. Model 4 has an adjusted R^2 of 0.114, which is higher than the adjusted R^2 of the regular OLS model. Three out of the five included city dummies are statistically significant. Relative to Model 1, the inclusion of the city dummies reduced the statistical significant of hand washing, water purification, assets, and debt.[3] This shows that city level variation exists in health outcomes. Negative surroundings, individual freedom and home ownership all stay significant at the 1% level even with the inclusion of the city dummies. The city a person lives in has a significant impact on their individual health outcomes, either positively or negatively. For example relative to Mumbai (city1), Delhi (city 2) and Kolkata (city

[3] Hyderabad (city 6) is not included due to collinearity. Mumbai and Bangalore are insignificant, thus not presented in the table.

3) have positive sign suggesting they have worse health outcomes, while Chennai (city 4) has a negative sign suggesting that individuals in that city report better self-reported health. This makes sense considering Delhi is known for having worse pollution statistic in comparison to other cities. Also from the raw data, that respondents in Delhi on average report that they have worse health in comparison to the other cities.

Model 5 shows the results of the multilevel random effects model. These are used when data is hierarchically organized and nested within categories. They allow you to see the impact that existing in these various "nests" impact the individual observations. Similar to the OLS model, individual freedoms and surroundings are significant at the 1% level for self-reported health and exhibit the same association as in previous models. Hand washing is significant at the 10% level and again shows the opposite of the expected result. R^2 for this model is not reported due to the fact that for multilevel models it is both difficult to obtain and often in accurate. In multilevel models, each residual has a different variance and thus a different R^2. There are common methods used to calculate R^2 for two level models, but it is much less common with three levels.

Table 6. Multilevel Regression Random Effects

	Self –Reported Health	
	Variance (SE)	% Total Variance
City	0.144 (0.050)	13.91%
Neighborhood	0.199 (0.024)	19.11%

City level and neighborhood level random effects exist for the self-reported health model and the days ill model.[4] The likelihood ratio test for both models has a $p > x^2 = 0$, rejecting the null hypothesis that there are no neighborhood level and city level random effects. For self-

[4] Days ill is reported in the appendix

reported health, around 14% of the total variance is accounted for at the neighborhood level and 19% is accounted for at the city level. While most variation still exists within neighborhoods, this result shows that there is variation both at the neighborhood and the city levels. The variation between neighborhoods is larger that between cities, which suggests that neighborhood within cities have higher degree of differences. Both the fixed effect results and the random effect results suggest that when designing interventions to improve health, one must consider the implications of location. It is important to design effective intervention that uplifts community level health and not only focus on individual level variables.

Results for days ill revealed similar results to self-reported health in terms of the neighborhood and city effects. Days ill still varied between cities and neighborhoods, but at a slightly smaller percentage. Significant variables are different between self-reported health and days ill. For the days ill model, the demographic variables are more significant. Household size, education, marriage age, hand washing and debt are all significant in the linear OLS model. Hand washing still exhibits opposite directionality. Older age at marriage is associated with better health outcomes. This shows similar results to Santhya *et. al* 2010 that the older a women is at marriage the better her self-reported health. Debt is also highly significant, with the largest magnitude. Women living in household that have outstanding debt also report higher numbers of days ill. Debt is less significant with a much smaller magnitude in the self-reported health models, which suggest debt is a more important factor for short-term health. Results for days ill models can be found in the Appendix. The similarity in results for the two variables suggests that short-term illness and more long-term health are both impacted by neighborhood effects. However, interventions looking to improve either health outcome would need to use different dependent variables as mechanism for change. While this paper does not focus on the difference

between short-term and long-term health, future exploration can be done to further answer this question. Other dependent variables such as days ill over a longer period, chronic illnesses, or maternal health outcomes could be used to explore this question in detail.

V. Conclusion

This study looks at the neighborhood level impact of health outcomes on women living in urban India. In the process, I find counter-intuitive results that suggest result that increasing frequency of hand washing resulted in worse self-reported health. This paper tries to explain these results using interaction terms and instrumental variables. While both interaction terms and instrumental variables are used in this study, neither provides expected results. This could be for a variety of reasons, include reporting errors or endogeneity issues. Further exploration is necessary to understand these results. The usage of soap while hand washing is not considered in this paper. Other variables not included in this model could also be interacting with hand washing, which could provide insight into the results. The instrumental variable used might also not clearly expose the true mechanism in which hand washing impacts an individual's health.

This paper finds that neighborhood and city level effects are present in urban India, while within neighborhood variation still accounts for around 2/3 of total variation. Any interventions looking at improving health outcomes in theses area should also consider this effect. Numerous interventions already do this, for example slum revitalization policy focuses on impacting an entire community. These programs might be more successful than those that just target individuals. This paper finds water purification, individual freedom, and assets are highly significant and increasing them leads to better self-reported health. Individual freedoms provides an interesting look at the idea that empowerment of women is an important aspect of their health. Recent programs such as microfinance loans targeted at women show examples of ways

interventions can engage women in the community and help improve health outcomes. Interventions focusing on individual women should also consider the community aspect of the intervention. Hand washing, negative surroundings, and home ownership are also highly significant. These three variables also have the largest magnitudes in comparison to other variables.

While study focuses on women's self-reported health in urban India, there are many more areas needing exploration. One of the major limitations of this study is the reliability of the dependent variables. Other measures of health could be used as dependent variables, such as incidences of illness, disabilities, physical tests, or blood pressure readings. Future studies could look at the differences between short-term and long-term health outcomes, in order to understand the mechanism that affect each distinct type of outcome. More longitudinal studies can help look at the way health outcomes change across time and with infrastructural changes in environments. These types of studies can help understand causality.

VI. References

Ackerson, L. K., Kawachi, I., Barbeau, E. M., & Subramanian, S. (2008). Geography of underweight and overweight among women in India: A multilevel analysis of 3204 neighborhoods in 26 states. *Economics & Human Biology, 6*(2), 264-280. doi:10.1016/j.ehb.2008.05.002

Ahmad, B., Ryan, V., Maziak, W., Pless-Mulloli, T., & White, M. (2013). The influence of neighbourhood formality status and socio-economic position on self-rated health among adult men and women: A multilevel, cross sectional, population study from Aleppo, Syria. *BMC Public Health, 13*(1). doi:10.1186/1471-2458-13-233

Aiello, A. E., PhD., Coulborn, R. M., B.S., Perez, V., M.S., & Larson, Elaine L,PhD., R.N. (2008). Effect of hand hygiene on infectious disease risk in the community setting: A meta-analysis. *American Journal of Public Health, 98*(8), 1372-81. Retrieved from http://proxy.lib.duke.edu/login?url=http://search.proquest.com/docview/215090126?accountid=10598

Arcaya, M. C., Tucker-Seeley, R. D., Kim, R., Schnake-Mahl, A., So, M., & Subramanian, S. (2016). Research on neighborhood effects on health in the United States: A systematic review of study characteristics. Social Science & Medicine, 168, 16-29. doi:10.1016/j.socscimed.2016.08.047

Barik, Debasis & Sonalde Desai. (2013). Determinants of Private Healthcare Utilization and Expenditure Patterns in India. Ch. 5 in IDFC India Infrastructure Report, The Road to Universal Health Coverage. Pp. 52-62.

Beck, S., Pulkki-Brännström, A.-M. & San Sebastián, M. (2015). Basic income – healthy outcome? Effects on health of an Indian basic income pilot project: a cluster randomised trial. Journal of Development Effectiveness, 7(1), 111-126.

Butala, N. , VanRooyen M., & Patel R. (2010) Improved health outcomes in urban slums through infrastructure upgrading. *Social Science & Medicine 71*(5), 935-940, doi.org/10.1016/j.socscimed.2010.05.037.

Chhapola, V., & Brar, R. (2015). Impact of an educational intervention on hand hygiene compliance and infection rate in a developing country neonatal intensive care unit. *International Journal Of Nursing Practice*, *21*(5), 486-492. doi:10.1111/ijn.12283

Clark, J., & Kearns, A. (2012). Housing Improvements, Perceived Housing Quality and Psychosocial Benefits from the Home. *Housing Studies*, *27*(7), 915-939.

Clayton, M., Liiares-Zegarra, J. M., & Wilson, J. O. (2015). Can Debt Affect Your Health? Cross Country Evidence on the Debt-Health Nexus. *SSRN Electronic Journal*. doi:10.2139/ssrn.2429025

Desai, Sonalde and Lijuan Wu. (2010). Structured Inequalities: Spatial Disparities in maternity Care in India, *Margin: A Journal of Applied Economics*, 4(3): 293-320.

Desai, Sonalde, and Reeve Vanneman. India Human Development Survey-II (IHDS-II), 2011-12. ICPSR36151-v5. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2016-08-01. http://doi.org/10.3886/ICPSR36151.v5

Dobe, M., Mandal, R. N., & Jha, A. (2013). Social Determinants of Good Hand-Washing Practice (GHP) Among Adolescents in a Rural Indian Community. *Family & Community Health, 36*(2), 172-177. doi:10.1097/fch.0b013e318282ac42

Ghosh, A., & Mukherji, A. (2014). Air Pollution and Respiratory Ailments among Children in Urban India: Exploring Causality. *Economic Development and Cultural Change, 63*(1), 191-222. doi:10.1086/677754

Ghosh, S. (2011, November 19). Catastrophic Payments and Impoverishment Due to Out-of-Pocket Health Spending: The Effects of Recent Health Sector Reforms in India. Economic & Political Weekly. doi:10.2139/ssrn.1658573

Ghosh, S., & Bose, S. (2012). Morbidity among urban children in India distinctions between slum and non-slum areas. *Internationales Asien Forum.International Quarterly for Asian Studies, 43*(1), 47-59,168,171.

Guo, B., Zhang, L., & Sherraden, M. (2009). Household assets and health in china: Evidence and policy implications. *China Journal of Social Work, 2*(1), 66-79. doi:http://dx.doi.org.proxy.lib.duke.edu/10.1080/17525090802692229

Krishnan, S. (2003, May 13). Gender, Caste, and Economic Inequalities and Marital Violence in Rural South India. *Health Care for Women International, 26*(1), 87-99. doi:10.1080/07399330490493368

Kulkarni, M. (2012). Social determinants of health: The role of neighbourhoods, psychological factors and health behaviours in predicting health outcomes for the urban poor in India. Journal of Health Psychology, 18(1), 96-109. doi:10.1177/1359105311430004

Laverack, G. (2006). Improving health outcomes through community empowerment: a review of the literature. *Journal of Health Population and Nutrition*, *24*(1), 113+. Retrieved from http://go.galegroup.com.proxy.lib.duke.edu/ps/i.do?p=AONE&sw=w&u=duke_perkins&v=2.1&it=r&id=GALE%7CA212107066&sid=summon&asid=952dfc62270d89b6bbdc1e80f11e82b3

Leventhal, T., & Brooks-Gunn, J. (2000). The neighborhoods they live in: The effects of neighborhood residence on child and adolescent outcomes. Psychological Bulletin, 126(2), 309-337. doi:10.1037//0033-2909.126.2.309

Mainassara, H. B., & Tohon, Z. (2014). Assessing the health impact of the following measures in schools in Maradi (Niger): Construction of latrines, clean water supply, establishment of hand washing stations, and health education. *Journal of Parasitology Research*, C1+. Retrieved from http://go.galegroup.com.proxy.lib.duke.edu/ps/i.do?p=AONE&sw=w&u=duke_perkins&v=2.1&it=r&id=GALE%7CA421568551&asid=e013b3ac854e9813471ebeb3567a180a

Menec, V. H., Shooshtari, S., Nowicki, S., & Fournier, S. (2010). Does the Relationship Between Neighborhood Socioeconomic Status and Health Outcomes Persist Into Very Old Age? A Population-Based Study. *Journal of Aging and Health, 22*(1), 27-47. doi:10.1177/0898264309349029

Nguyen, N. e. (2012). Intestinal parasitic infection and nutritional status among school children in Angolela, Ethiopia. J Prev Med Hyg. , 53 (3), 157-64

Rosenblum, Daniel, 2015, "Unintended consequences of women's inheritance rights on female mortality in India," *Economic Development & Cultural Change* **63**(2): 223-248.

Sallis, J. F., Saelens, B. E., Frank, L. D., Conway, T. L., Slymen, D. J., Cain, K. L., Kerr, J. (2009). Neighborhood built environment and income: Examining multiple health outcomes. *Social Science & Medicine, 68*(7), 1285-1293. doi:10.1016/j.socscimed.2009.01.017

Santhya, K., Ram, U., Acharya, R., Jejeebhoy, S. J., Ram, F., & Singh, A. (2010, September). Associations Between Early Marriage and Young Women's Marital and Reproductive

Health Outcomes: Evidence from India. International Perspectives on Sexual and Reproductive Health, 36(03), 132-139. doi:10.1363/3613210

Spears, D. (2012). How much international variation in child height can sanitation explain?

Sood, N., Bendavid, E., Mukherji, A., Wagner, Z., Nagpal, S., & Mullen, P. (2014). Government health insurance for people below poverty line in india: Quasi-experimental evaluation of insurance and health outcomes. *BMJ : British Medical Journal (Online), 349* doi:http://dx.doi.org/10.1136/bmj.g5114

Subbaraman, R., O'brien, J., Shitole, T., Shitole, S., Sawant, K., Bloom, D. E., & Patil-Deshmukh, A. (2012). Off the map: The health and social implications of being a non-notified slum in India. *Environment and Urbanization, 24*(2), 643-663. doi:10.1177/0956247812456356

Sweet, E., Nandi, A., Adam, E. K., & Mcdade, T. W. (2013). The high price of debt: Household financial debt and its impact on mental and physical health. *Social Science & Medicine, 91*, 94-100. doi:10.1016/j.socscimed.2013.05.009

Vargese, S., Mathew, P., & Mathew, E. (2013). Utilization of public health services in a rural area and an urban slum in Western Maharashtra, India. *Int J Med Sci Public Health International Journal of Medical Science and Public Health, 2*(3), 632. doi:10.5455/ijmsph.2013.220420135

Wankhade, K. (2015). Urban sanitation in India: Key shifts in the national policy frame. Environment and Urbanization, 27(2), 555-572. doi:10.1177/0956247814567058

Appendix I. Independent Variables

Category	Name	Value Explanation
Health	Daily Disabilities	1 = disability 0 = no disability
	Hand washing	1 = never wash 4 = always wash
	Water Purification	1 = never purify 4 = always purify
Empowerment	Community Crime	1 = crime happens 0 = no crime
	Own Health	0 = other have more say 1 = respondent most say
	Individual Freedom	0 = not much freedom 6 = independent decisions
	Surroundings	1 = feces or stagnant water around house 0= no feces or stagnant water around house
Financial Security	Debt	1 = debt 0 = no debt
	Life insurance	0 = no life insurance 1 = life insurance
	Home owner	0 = rent 1 = own

Constructed Variables

Community Crime = an index of 3 question: theft, break-in or attack in neighborhood in last 6 months

Individual Freedom = an index of 6 questions: respondent decides number of children, can go short distances alone, can visit friend/relative alone, has cash-in hand for household expenses, respondent decides purchasing of expensive items, and respondents name on homeownership papers)

Assets = a combination of 30 household assets, constructed within the dataset

Surroundings = a combination of 2 questions, existence of feces or stagnate water in front of household

Life Insurance = a combination of 2 questions, household has private or government life insurance

Cross Correlations

	age	person~HE	Education	marria~e	handwash	waterp~y	commun~e	ownhea~h	indivi~m	ASSETS	hhsurr~d	INCOME	debt	lifeinsura~
age	1													
personsinHH	-0.0516	1												
education	0.0803	0.0146	1											
marriageage	-0.0381	-0.1021	0.4467	1										
handwash	0.0432	0.0009	0.1367	0.0877	1									
waterpurify	0.1003	0.003	0.2568	0.1627	0.0393	1								
communityc~e	0.0139	0.0116	0.0154	-0.0069	0.0903	0.0216	1							
ownhealth	0.0927	-0.0621	-0.0015	-0.0401	0.0504	0.0155	0.0061	1						
individual~m	0.1631	-0.0683	0.1571	0.0201	0.0251	0.008	-0.0304	0.1182	1					
ASSETS	0.14	0.1103	**0.6069**	0.347	0.1251	0.3104	0.0352	-0.0276	0.1947	1				
hhsurround	-0.0397	0.0625	-0.2174	-0.1476	-0.0243	-0.0843	-0.0067	-0.0228	-0.032	-0.2268	1			
INCOME	0.0978	0.1807	0.3927	0.2311	0.0581	0.2288	0.0242	-0.0013	0.0756	0.4954	-0.1237	1		
debt	-0.0684	0.0269	-0.1953	-0.1716	-0.1138	-0.0708	0.0155	0.0256	-0.104	-0.2541	0.1039	-0.1208	1	
lifeinsura~e	0.1197	-0.0395	0.318	0.1843	0.1106	0.1628	0.0754	-0.0443	0.0802	0.3744	-0.0959	0.2716	-0.1068	
homeowner	0.0587	0.049	-0.0277	-0.0253	0.0353	-0.0246	0.0159	-0.0168	0.0166	-0.1017	0.0545	-0.0534	-0.0009	-0.08

Appendix II. Log Models

Category	Variables	Log SRH	Log Days Ill
Demographic	Log(Age)	12.045	17.173
	Log(Age)^2	-3.526	-4.932
	Log(Age)^3	0.345	0.471
	Log(HH size)	-0.005	-0.145
	Log(Education)	-0.007	-0.038
	Caste/Religion	Insignif.	Insignif.
	Log(Marriage Age)	-0.036	-0.285
Health	Daily Disabilities	0.087	0.332
	Log(Hand washing)	0.146	0.236
	Log(Water Purification)	-0.041	0.075
	Log(Miscarriages)	0.119	0.180
Empowerment	Community Crime	0.017	0.055
	Own Health	0.028	0.022
	Log(Individual Freedom)	-0.133	-0.004
HH Assets	Log(Assets)	-0.093	-0.009
	Surroundings	0.085	0.017
Finances	Log(Income)	-0.022	-0.075
	Debt	0.020	0.104
	Life insurance	0.005	0.036
	Home owner	0.067	-0.020
	Constant	-12.713	-18.1919
	Adjusted R^2	0.091	0.054
	N	2,821	2,821

Appendix III. Days Ill Results

Category	Variables	Model (1) OLS	Model (2) Interaction	Model (3) IV	Model (4) Fixed E.	Model (5) Random E.
Demographic	Age	0.120 (0.241)	0.119 (0.241)	-0.097 (1.727)	0.194 (0.239)	0.182 (0.238)
	Age^2	-0.003 (0.006)	-0.003 (0.006)	0.012 (0.063)	-0.005 (0.006)	-0.005 (0.006)
	Age^3	0.000 (0.000)	0.000 (0.000)	0.000 (0.001)	0.000 (0.000)	0.000 (0.000)
	HH size	-0.087*** (0.033)	-0.087*** (0.033)	-0.112** (0.026)**	-0.110*** (0.034)	-0.110*** (0.033)
	Education	-0.029* (0.016)	-0.029* (0.016)	-0.051 (0.026)	-0.016 (0.016)	-0.015 (0.016)
	Caste/Religion	Insignif.	Insignif.	Insignif.	Insignif.	Insignif.
	Marriage Age	-0.054*** (0.019)	-0.054*** (0.019)	-0.044 (0.030)	-0.038** (0.019)	-0.042** (0.019)
Health Practices	Hand washing	0.387*** (0.113)	0.520** (0.260)		0.390 (0.269)	0.250** (0.123)
	Hand washing IV			0.598 (0.588)		
	Water Purification	0.108 (0.053)	0.383 (0.487)	0.098 (0.078)	0.435 (0.495)	0.127** (0.055)
	Interaction		-0.071 (0.126)			
Empowerment	Community Crime	0.174 (0.237)	0.178 (0.237)	-0.314 (0.375)	0.286 (0.238)	0.269 (0.237)
	Own Health	0.060 (0.160)	0.062 (0.160)	0.056 (0.253)	0.116 (0.163)	0.080 (0.162)
	Individual Freedom	0.030 (0.054)	0.031 (0.054)	-0.048 (0.076)	-0.017 (0.055)	-0.001 (0.055)
HH Assets	Assets	-0.018 (0.016)	-0.019 (0.016)	0.031 (0.024)	-0.012 (0.018)	-0.019 (0.017)
	Surroundings	0.030 (0.156)	0.024 (0.156)	0.076 (0.217)	-0.078 (0.156)	-0.055 (0.156)
Financials	Income	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000** (0.000)	0.000 0.000
	Debt	0.525*** (0.123)	0.526*** (0.123)	0.719*** (0.185)	0.441*** (0.129)	0.439*** (0.128)
	Life insurance	0113 (0.131)	0.111 (0.131)	0.173 (0.190)	0.122 (0.134)	0.141 (0.133)
	Home owner	0.065 (0.139)	0.069 (0.139)	-0.242 (0.194)	0.045 (0.140)	-0.003 (0.141)
City Dummies	City 1				-1.186*** (0.265)	
	City 3				-0.433* (0.247)	
	City 4				-1.528*** (0.303)	
	City 5				-0.975*** (0.284)	
	Constant	-0.834 (3.567)	-1.339 (3.678)	-1.826 (15.876)	-1.094 (0.284)	-1.265 (02.535)
	Adjusted R^2	0.027	0.026	.	0.0414	.
	N	2,844	2,844	1,269	2,844	2,844

	SRH		Days Ill	
	Variance (SE)	% Total Variance	Variance (SE)	% Total Variance
City	0.144 (0.050)	13.91%	0.394 (0.154)	10.32%
Neighborhood	0.199 (0.024)	19.11%	0.368 (0.367)	9.62%