Isolation and Subjective Welfare: Evidence from South Asia*

Marcel Fafchamps — Oxford University†

Forhad Shilpi — The World Bank‡

March 2008

Abstract

Using detailed geographical and household survey data from Nepal, this article investigates the relationship between isolation and subjective welfare. We examine how distance to markets and proximity to large urban centers are associated with responses to questions about income and consumption adequacy. Results show that isolation is associated with a significant reduction in subjective assessments of income and consumption adequacy, even after controlling for consumption expenditures and other factors. The reduction in subjective welfare associated with isolation is much larger for households that are already relatively close to markets. These findings suggest that welfare assessments based on monetary income and consumption may seriously underestimate the subjective welfare cost of isolation.

JEL codes: D60, I31, R20

Keywords: geographical isolation; consumption adequacy; subjective well-being

---

*We benefitted from comments from Ravi Kanbur, Mukesh Eswaran, John Hoddinott, John Pender, Martin Ravallion, and participants to a seminar at Cornell and IFPRI. We also received excellent comments on an earlier version of this paper from participants to the Cornell/LSE/MIT Conference on Behavioral Economics, Public Economics and Development Economics; participants to the Global Poverty Research Group workshop held in June 2004; and from participants to seminars at LSE, CERDI, the University of British Columbia, Bath University, and the World Bank. We would like to thank Migiwa Tanaka for excellent research assistance. The support of the World Bank and of the Economic and Social Research Council (UK) is gratefully acknowledged. The work was part of the programme of the ESRC Global Poverty Research Group.

†Department of Economics, University of Oxford, Manor Road, Oxford OX1 3UQ. Email: marcel.fafchamps@economics.ox.ac.uk.

‡DECRG, The World Bank, 1818 H Street N.W., Washington DC 20488 USA..

1. Introduction

While much has been written on the relationship between geographical location and objective measures of consumption and welfare (e.g. Elbers, Lanjouw & Lanjouw 2003, Jalan & Ravallion 2002, Ravallion & Jalan 1996, Ravallion & Jalan 1999, Ravallion & Wodon 1999), little is know on how isolation affects subjective welfare. The traditional literature on labor migrations has often assumed that rural dwellers prefer living in the countryside and would have to be compensated for migrating to town. This assumption, for instance, underlies original contributions by Lewis (1954) and Harris & Todaro (1970). More recently, Murphy, Shleifer & Vishny (1989) make a similar assumption regarding wage work. Little hard evidence however exists on the utility cost or benefit of rural living.

This paper revisits the question of the relationship between isolation and subjective welfare and estimate the welfare cost of geographic isolation. To this effect, we use answers to subjective questions about consumption and income adequacy to test whether utility is equalized across space and, if it is not, whether utility is higher or lower in isolated areas. This approach enables us to investigate in a direct and straightforward manner the question of the relationship between isolation and utility without requiring any assumption about spatial mobility.

The starting point of our empirical specification is a standard utility maximization model in which isolation is related to utility through its effect on incomes and prices, on the availability of goods and services, and on public goods and externalities. For our empirical investigation, we use a large-scale living standard measurement survey, the Nepal Living Standard Survey (NLSS) of 1995/96. Nepal is the perfect country to study isolation because so much of the country remains inaccessible by road. The NLSS includes a number of questions on subjective consumption and income adequacy. The head of each surveyed household was asked to rank the household's total income as 'not adequate', 'adequate' or 'more than adequate'. Similar questions were asked about five consumption categories, namely food, clothing, housing, schooling, and health care. We investigate whether responses to these questions vary systematically with distance to markets and cities.[1]

Our econometric investigation leads to a robust finding: isolation is associated with lower subjective welfare. This result obtains after we control for consumption expenditures, suggesting that the relationship between isolation and welfare is not only due to lower monetary consumption. Controlling for household mobility and adding various controls leaves results unchanged. We quantify the difference in subjective welfare associated with isolation and find it to be large, particularly for housing, schooling and health care. Surprisingly, the reduction in subjective welfare associated with isolation is largest for households already close to markets. These results should be interpreted as indicative of a strong empirical relationship between geographical isolation and subjective welfare. Better data is needed to ascertain the causal effect of geographical isolation on welfare.

The paper is organized as follows. The conceptual framework is discussed in Section 2. Section 3 describes the data and its main characteristics. Econometric estimation results are presented in Section 4 while in Section 5 we quantify the reduction in subjective welfare associated with geographical isolation. Section 6 discusses the results and Section 7 concludes the paper.

2. Conceptual framework

For the sake of this paper, let us define isolation as distance from urban centers: a household located at a large distance d from the nearest urban center is deemed to be more isolated than a household located closer. We are interested in the relationship between welfare and local

[1] Other surveys have also asked what is typically referred to as the subjective well-being question, namely, 'Do you feel generally happy with life?'. Unfortunately, this question was not asked in the NLSS.

characteristics. To capture this idea, let individuals derive utility V from total consumption expenditures X and from public amenities A:

$$V_{jk} = V(X_{jk}, A_k)$$

where j denotes the individual and k the location. To keep the presentation simple, other factors such as prices, product variety, etc, are ignored for now. We introduce them later. We ignore savings, so that income equals consumption.

Individuals prefer the location k where their utility V_{jk} is highest. Whether they can relocate or not depends on the functioning of the labor market. We first discuss the case in which workers locate freely and costlessly. We then examine the case where workers are immobile or move at a cost.

2.1. The cost of isolation

Assume for a moment that A is the same across locations. If individuals can move at no cost, arbitrage implies that utility – and hence income – are equalized across locations. To generate different levels of income and utility across locations, let us follow Roy (1951), as modified by Dahl (2002), and assume that workers differ in ability ε_j so that some individuals have higher marginal productivity. In a competitive labor market with free movement of labor, workers are paid their marginal product. We thus have $X_{jk} = X_k(\varepsilon_j)$ with $\partial X_k(\varepsilon)/\partial\varepsilon > 0$.

Now assume that, for technological reasons, jobs that require a high ability are located in or near cities, i.e., that $\partial^2 X_k(\varepsilon)/\partial\varepsilon\partial d_k < 0$.[2] With these assumptions, average welfare is higher

[2] This is not an unreasonable assumption. Fafchamps & Shilpi (2005), for instance, have shown that there are larger firms and more job specialization in and around cities. We also know from the work of Jacoby (2000) that, in Nepal, land located far from markets has a lower value and yields a lower income. This is undoubtedly due to lower average prices for agricultural output and less emphasis on commercial farming, an issue that is revisited by Fafchamps & Shilpi (2003). All these factors probably generates higher returns to education and to entrepreneurial ability in urban centers.

in cities, i.e., $\partial E[V|d]/\partial d < 0$. This is because, in equilibrium, high ability workers locate in urban centers where wages are higher. If we assume that ability has no direct effect on utility, welfare differences across locations and individuals are entirely driven by differences in income. Once we condition on income, we should observe no systematic relationship between utility and urban proximity, i.e., $\partial E[V|X,d]/\partial d = 0$.

Now assume that market towns and other urban centers have better amenities – higher A. This could be because it is less costly to provide public services to a concentrated population. Since workers locate freely, arbitrage implies that:

$$V(X_k(\varepsilon_j), A_k) = V(X_m(\varepsilon_j), A_m)$$

where k and m denote two different locations. It follows that $X_{jk} < X_{jm}$ if $A_k > A_m$: for a given ability ε_j, workers in high A areas receive a lower wage than workers in low A areas (e.g. Rosen 1979, Roback 1982).[3] By assumption A is a decreasing function of distance from cities. It follows that $\partial X_k(\varepsilon_j)/\partial d_k > 0$.

This generates testable predictions. If we compare two workers earning the same income but living in locations with different levels of amenities A, it must be that the worker in the better location has higher ability, and thus higher utility. There is a negative relationship between utility and the lack of amenities – proxied by distance – after we control for income. It is possible to measure the implicit value of amenities by comparing wages of similar ability workers across locations with different levels of amenities. This is the approach adopted, for instance, by Rosen (1979) and Roback (1982).

Alternatively, suppose we do not observe ability but we observe a strictly increasing monotonic

[3] Of course, since ability is higher in urban areas, the average income across workers of different abilities is higher in urban areas and $\partial E[X_k(\varepsilon_j)]/\partial d_k < 0$ where the expectation is taken over all abilities ε.

transformation $W_{ij} = g(V_{jk})$. Strict monotonicity of $g(.)$ implies that $g(V_{jk}) = g(V_{jm}) \Leftrightarrow V_{jk} = V_{jm}$. Further suppose that we do not observe A_k but $A(d_k) = Ad_k^{-\alpha_2}$. We can estimate the implicit utility cost of isolation by regressing W_{jk} on X_{jk} and d_k:

$$W_{jk} = \alpha_0 + \alpha_1 \log X_{jk} - \alpha_2 \log d_k + u_{jk} \tag{2.1}$$

Differences in utility across location reflect differences in ability, but by comparing individuals with the same utility and different consumption X_{jk} we identify the effect of distance d_k on utility. Indeed, controlling for X_{jk}, utility V_{jk} falls with distance from urban centers as amenities get worse. This simple observation constitutes the basis of our testing strategy.

To compute the equivalent variation of isolation, let C_{km} denote the proportion of income that makes individual j indifferent between distance d_k and distance d_m. Equivalent variation $C_{km}X_{jk}$ is the value that satisfies:

$$\alpha_0 + \alpha_1 \log X_{jk} - \alpha_2 \log d_k + u_j = \alpha_0 + \alpha_1 \log(X_{jk} - C_{km}X_{jk}) - \alpha_2 \log d_m + u_j$$

which after some straightforward algebra yields:

$$C_{km} = 1 - e^{\frac{\alpha_2}{\alpha_1}(\log d_k - \log d_m)} \tag{2.2}$$

In case workers are immobile, the method proposed by Rosen (1979) and Roback (1982) breaks down. Since utility is not equalized across locations, it cannot be assumed that better amenities are compensated by lower wages, and wage differences across locations for workers of the same ability cannot be interpreted as the hedonistic price of better amenities.

The utility approach still works, however. It also works if workers can only move at a cost, or if some individuals can move and others cannot – for instance because of credit constraints or of discrimination in the labor market. This feature is particularly appealing given the kind of data we have. Labor markets in Nepal are not as fluid as they are in the US. According to (Dahl 2002), over 30% of US employees work in a state other than their birth state. In contrast, in Nepal, a country of 20 million people, more than 80% of household heads reside and work in their birth *village*. For the Roback approach to work, it is not required that all workers be mobile – only that, in each ability category and each location, some workers be mobile so that, at the margin, the arbitrage argument works. But with so many workers immobile, it is likely that arbitrage fails for at least some locations and some ability categories, therefore invalidating the Roback test. Furthermore, the overwhelming majority of Nepalese people are self-employed. Their income depends on dimensions of individual ability that are difficult to measure – such as experience, familiarity with local conditions, and entrepreneurial spirit. It is therefore unlikely that we would be able to control for ability sufficiently well to measure the value of amenities using the Roback approach.[4]

2.2. Multiple subjective satisfaction indicators

So far we have discussed the case of a single utility indicator. Now suppose that we have subjective satisfaction indicators for consumption subsets c_h such as food or clothing. To integrate these indicators into the analysis, we decompose consumption into H subsets and we assume that utility is (approximately) Cobb-Douglas with respect to these subsets. We start by ignoring

[4] Dahl (2002) proposes a way of dealing with selection on unobserved ability. This solution requires not only a number of additional assumptions but also massive amounts of data to compute migration probabilities between each location. Unfortunately we do not have sufficient data to compute such transition matrices for Nepal.

amenities A. Dropping jk subscripts for easier reading, we have:

$$V = \sum_{h=1}^{H} \omega_h \log c_h$$

where the ω_h's are consumption shares, with $\sum \omega_h = 1$. Let V^h be the sub-utility obtained from the consumption of good h:

$$V^h = \log c_h$$

If the consumer chooses consumption optimally, we have:

$$\begin{aligned} V &= \sum_{h=1}^{H} \omega_h \log \frac{\omega_h X}{p_h} \\ &= a + \log X - \log P \end{aligned}$$

where a is a constant and P is a price index defined as $P = \prod_h p_h^{\omega_h}$. Similarly we can write:

$$\begin{aligned} V^h &= \log \frac{\omega_h X}{p_h} \\ &= b + \log X - \log p_h \end{aligned}$$

where $b = \log \omega_h$ is a constant.

To introduce geographical isolation, suppose that $p_h = p d^{\lambda_h}$ where parameter λ_h captures differences in amenities and in transport costs across consumption subsets. Taking logs we get:

$$V^h = b' + \log X - \lambda_h \log d \tag{2.3}$$

By comparing λ_h across consumption subsets, we can infer which consumption subsets are most sensitive to isolation.

In practice, we do not observe V^h directly but a proxy W^h, namely the likelihood of answering 'inadequate', 'adequate' or 'more than adequate' to a consumption adequacy question for subset h. Regression estimation yields:

$$\begin{aligned} W^h &= g_h(V^h) = \alpha_0^h + \alpha_1^h \log X - \alpha_2^h \log d \\ V^h &= g_h^{-1}(\alpha_0^h + \alpha_1^h \log X - \alpha_2^h \log d) \end{aligned} \tag{2.4}$$

Totally differentiating (2.4) and (2.3) and setting them equal, we get:

$$\begin{aligned} \frac{\partial V^h}{\partial \log X} &= \frac{\partial g_h^{-1}}{\partial V^h}\alpha_1^h = 1 \\ \frac{\partial V^h}{\partial \log d} &= -\frac{\partial g_h^{-1}}{\partial V^h}\alpha_2^h = \lambda_h \end{aligned}$$

from which it follows that:

$$\lambda_h \approx \frac{\alpha_2^h}{\alpha_1^h} \tag{2.5}$$

where the approximation comes from the fact that we are averaging over observations. The equivalent variation C_{km}^h of isolation can be calculated for each consumption subset using λ_h from equation (2.2).[5]

So far we have assumed that consumers face no quantity rationing. This may be a reasonable

[5] Product diversity can also be introduced to the model as follows. Assume that to each consumption subset h there corresponds an aggregation function of the form:

$$c_h = \left(\int_0^{N_h} c(s)^{\varphi} ds \right)^{\frac{1}{\varphi}}$$

where s denotes a continuum of goods and N_h determines the range of goods available. If the prices of all goods are identical, we have $c_h = cN_h^{\frac{1}{\varphi}}$. Inserting into the utility function, we obtain an extra term:

$$V^h = b + \log X - \log p_h + \frac{1}{\varphi} \log N_h$$

This shows that utility decreases with a fall in variety N_h, resulting for instance from isolation. To simplify notation, we omit variety N_h from the rest of the presentation.

assumption for many goods but it is inadequate for public goods such as law enforcement or clean air. It is also problematic for goods that are publicly provided at a subsidized price, such as health care. For these amenities, quantitative rationing arises if individuals are unable to purchase what they wish to consume at the subsidized price.

To illustrate this case, let us partition goods into rationed r and unrationed u.[6] Many amenities fall into the rationed category. Utility is written:

$$V = \sum_{u=1}^{U} \omega_u \log c_u + \sum_{r=1}^{R} \omega_r \log c_r \tag{2.6}$$

where c_r is regarded as exogenously determined – it is not a choice parameter of the household. Define $X_u = X - \sum p_r c_r$. Utility maximization over the unrationed goods yields the familiar demand functions $\frac{\omega_u X_u}{p_u}$ and the indirect utility function can be written:

$$V = b + \log X_u - \log P_u + \sum_{r=1}^{R} \omega_r \log c_r \tag{2.7}$$

As we can see, utility now depends also on the consumed quantity of public goods and rationed public services. Suppose that c_r varies with isolation d. When we regress V on X_u and distance d, the coefficient of d also captures difference in $\sum_{r=1}^{R} \omega_r \log c_r$, the value of which is included in the equivalent variation of isolation C_{km}.

When we look at specific consumption subsets, however, the utility derived from c_r drops

[6] By rationing we mean that the consumer is off his demand curve. If limited supply of a good, say l, results in a high price in location k, we do not regard this as rationing: since the consumer is on his demand curve, the formulas for the unrationed goods apply.

out:

$$V^u = \log c_u$$
$$= b + \log X_u - \log p_u \quad (2.8)$$

From this difference between (2.7) and (2.8), it follows that comparing C_{km} for total consumption with C^h_{km} for specific consumption subsets provides information about the isolation cost associated with public goods.[7]

For rationed goods c_r we have:

$$V^r = \log c_r$$

Suppose that $c_r \approx cd^{-\lambda_r}$. It follows that $V^r = \log c - \lambda_r \log d$. We can thus estimate a regression of the form:[8]

$$W^r = g_r(V^r) = \alpha^r_0 - \alpha^r_2 \log d$$

This suggests a way of identifying rationed goods: for them, total expenditures X_u do not enter the regression.

Since expenditures do not enter the equation for W^r, we cannot estimate λ_r using (2.5) because we do not have α^r_1. We may, however, obtain an order of magnitude for λ_r if we are willing to make some strong assumptions regarding the way subjective adequacy questions were answered. Suppose we are willing to assume that $g_u(.) \approx g_r(.)$. This equivalent to saying that individuals answer adequacy questions about one subset in a way that is commensurate to the

[7] It is imporant to recognize that this comparison holds strictly only for Cobb-Douglas preferences. For more general preference functions, the consumption of amenities c_r may affect satisfaction derived from unrationed goods because of complementarities between subsets. We revisit this issue below.

[8] It is conceivable that consumption is rationed for certain consumers but not others. we will be estimating a model that is a mixture of the rationed and unrationed case. Although we do not discuss this case explicitly here, it is intuitively clear that, as a result of attenuation bias, the coefficient of total expenditures will be smaller.

contribution of the subset to total utility. If this were the case, α_1^u would be the same for all unrationed goods and we could also use it to normalize the welfare cost of isolation for rationed goods:

$$\lambda_r \approx \frac{\alpha_2^r}{\alpha_1^u} \tag{2.9}$$

If we have different unrationed goods u, we will have different estimates of λ_r, one for each unrationed good α_1^u. If estimates of α_1^u for unrationed goods are relatively similar, we may hope that – should good r be unrationed – its α_1^r would fall within the same range. With these strong assumptions, we can 'bracket' λ_r and, by extension, C_{km}^r.

3. The data

The data we use come from the Nepalese Living Standard Measurement Survey (LSMS) of 1995/96. The survey drew a nationally representative sample of 3373 urban and rural households spread among 274 villages or 'wards'. Between 16 and 20 households were interviewed in each ward. As with other LSMS surveys, data coverage of NLSS 1995/96 is quite comprehensive.

The survey includes a series of questions on the adequacy of consumption level enjoyed by the household. The household head was first asked the following question: "Concerning [your family's food consumption over the past one month], which of the following is true? It was less than adequate for your family's needs [1], it was just adequate for your family's needs [2], it was more than adequate for your family' needs [3]." The household head was then asked five other similar questions in which the part in brackets is replaced by: [your family's housing], [your family's clothing], [the health care your family gets], [your children's schooling], and [your family's total income over the past one month], respectively.

Responses to these questions are summarized in Table 1. The overall dissatisfaction of household heads is quite striking. About 69 percent of household heads feel they have less than adequate income. Even for food consumption, which receives the best adequacy rating of the six questions, 47 percent of the household heads report it to be inadequate relative to needs. Only a small proportion of households report their income or consumption to be more than adequate. Although disturbing, these figures are consistent with more objectively measured welfare: at the time of the LSMS survey, 42% of the Nepalese population was estimated to be below the poverty line (World Bank, 1999).

Household characteristics are summarized in Table 2. The Nepal survey contains detailed information about travel time to a number of different facilities. Given that Nepal is a very mountainous country, distance in Km is not a relevant measure for most of the country; travel time is a more accurate measure of isolation in this case. We see that, on average, surveyed households live on average more than two hours of travel time from a market, the maximum value being 40 hours.[9] The median is around 1 hour. Distance to local markets is the first isolation measure d_{jk} that we use in our empirical analysis. Given the nature of the terrain and the spatial dispersion of households, d_{jk} varies between individuals within the same ward. Travel times to the nearest school and health facility are much shorter: on average households are located around 20 minutes from the nearest school and one hour from the nearest health facility. The quality of schools and health facilities varies widely across locations, however.

Average total annual consumption (non-durables and durables) is reported in US$.[10] The total value of assets is reported next. This includes land, livestock, agricultural equipment, and financial assets. As is customary, wealth distribution is quite unequal (high standard deviation) and highly skewed, with the median representing around one fourth of the mean. Parental

[9] Our measure of isolation, 'distance to markets' is computed as the average of the travel time to five different types of markets, namely market centers, *hat/bazar*, *krishi* center, cooperative center and local shops. Taking any single one of them leads to the loss of many observations. This information is recorded independently for each household.

[10] Using the exchange rate of 56.8 Rupees to the dollar which prevailed at the time of the survey. For reference, in the regression analysis we use (logged) Rupees.

background variables are reported as well, such as land inherited by the household, education level of the father of the household head, and whether the head's father was employed in a non-farm occupation. Later on we use these variables to predict migration out of one's birth ward. In 1996, towards the end of the NLSS survey, a Maoist insurgency began to take root in rural Nepal. Since the insurgency initially limited itself to attacking a few police stations, it had a minimal direct impact on the welfare of survey respondents. But it may have affected their expectations regarding the future, raising the possibility of omitted variable bias. At the bottom on Table 2 we report insurgency incidence figures based on a June 2000 classification of the Nepalese police. Some 12.5% of the surveyed households resided in areas that were seriously affected by the insurgency between 1996 and 2000. These districts tend to be far from urban centers.

Ward-level variables are presented in Table 3. Using detailed information on the road distance between each ward and each of 34 towns and cities compiled by Fafchamps & Shilpi (2003), we construct a variable that represent the total urban population P_k living within 2 hours of travel distance from the ward. Population figures come from the 1991 census. This is our second isolation measure. Our third isolation measure is population density in the district. Other things being equal, we expect people in low population density districts to live further apart from each other, thereby raising delivery costs for private goods as well as public services.

The survey did not collect extensive price data. There is information on house rental prices, mostly on a self-assessed basis. In the next section we combine this information with house quality data to estimate a district-specific house price premium. This premium is thought to capture locational advantages reflected in housing prices. We have information on rice prices at the household level, from which we compute a ward-level median.[11] We use the wage rate in the

[11] Household-level prices capture differences in quality, quantity and convenience between households facing the same market. It is more reasonable to assume that all households residing in the same ward face the same prices.

ward as an additional measure of the cost of living. We compute the median wage rate in the ward from responses of individual household members. Nearly all wage employment recorded in the survey is for low skilled manual work in farm and non-farm work. We report Gini coefficients for consumption per capita computed for each ward. We use it as control. It is indeed thought that inequality affects subjective well-being negatively. If inequality is stronger in urban areas, this could generate an omitted variable bias.

To capture the impact of product variety N, we construct for each ward indices of variety for food, non-food, and durables. For each household, the survey collected consumption expenditure information on 67 separate food items, 58 non-food items and 16 durable goods items. Based on this information, we compute the total expenditures s_i by all surveyed households in a ward on item i. Not all items are consumed in any given ward. For instance, of the 67 food items listed in the questionnaire, some wards consume 63 items while others only consume 33. Based on this information, we compute, for each of our three groupings J (food, non-food, and durables), a Herfindahl concentration index defined as:[12]

$$N^J = \frac{\sum_{i \in J} s_i^2}{[\sum_{i \in J} s_i]^2}$$

This index gives a rough idea of what is available for sale in the ward: the higher its value, the more concentrated spending is on a small number of categories, and the less diversified ward consumption is.[13] N^J does not, however, measure product diversity within each sub-category and is thus an imperfect measure of variety. Index values reported in Table 3 show

Because of the presence of outliers (probably due to measurement error), we use the median price in the ward rather than the mean.

[12]To the extent that richer consumers buy a greater variety of products, our variety indices are correlated with consumption expenditures. This is not a cause for concern, however, since we control for consumption expenditures directly in the regression analysis.

[13]$N^J = 1$ corresponds to complete concentration in a single item while $N^J = 1/J$ corresponds to equal expenditure shares for all J items in a category.

more concentration in durable expenditures and less in foodstuffs.

Two sets of dummies control for climatic and economic factors: ecological belt dummies and regional dummies. Ecological belt dummies divide the country into three North-South zones based on elevation. The mountain zone is the part of the country located at 4000 meters (12000 feet) of elevation and above. The Terai is the narrow plain bordering India. The Hills is the intermediate zone where much of the Nepalese population lives. Regional dummies capture an East-West division of the country. The Central region is where the capital Katmandu is located. We also use average rainfall and rainfall variability between years as additional proxies for local agro-climatic conditions.

4. Econometric results

We now turn to the econometric analysis. Responses to subjective adequacy questions – coded from 1 to 3 – are the dependent variables used in our analysis. There are six dependent variables: satisfaction with food, clothing, housing, schooling, health care, and total income. Satisfaction with total income should, in principle, combine all the effects of isolation and can be taken as proxy for W while answers to questions about specific consumption groups proxy for W^h. It is, however, conceivable that respondents regard monetary issues as separate from problems of product variety (N) and access to amenities (A). Someone may, for instance, answer that his income is adequate but complain that he cannot buy the clothing or health care he desires because it is not available locally. If, for most respondents, product variety and access to public services are conceptually distinct from the magnitude of monetary income, answers to the income adequacy question may fail to include the effects of N and A – and thus be less sensitive to isolation.

4.1. Non-parametric analysis

We begin with non-parametric univariate regressions of answers to income and consumption adequacy questions on the log of distance to markets. The purpose of the exercise is to document the existence of a strong correlation between the two and to illustrate that the relationship between adequacy responses and $\log d_{jk}$ is approximately linear.

Results are presented in Figures 1a to 1f, using an Epanechnikov kernel with moderate smoothing. The 95% confidence interval is also reported to facilitate inference.[14] It is immediately apparent that subjective consumption adequacy falls dramatically and significantly with distance from markets.[15] The relationship between $\log d_{jk}$ and subjective adequacy is monotonic and basically linear, except at high market distances for which the small number of observations does not allow precise estimation. This means that subjective adequacy falls rapidly at short distances, before tapering off. In the rest of the analysis we use $\log d_{jk}$ as regressor.

As explained in the conceptual section, the relationship depicted in Figures 1a to 1f could be the result of selection by ability. To investigate this possibility, we perform a non-parametric regression of consumption expenditures on distance. For the regression to be meaningful, we need to control for differences in household size and composition. One approach would be to divide total expenditures by the number of household members, possibly weighted by gender and age, yielding consumption per adult equivalent. But doing so may bias results due to economies of size in household production (e.g. Deaton & Paxson 1998, Fafchamps & Quisumbing 2003). To avoid such bias, we use a semi-parametric regression of the form:

$$\log X_{jk} = \beta_1 B_{jk} + \beta_2 \log p_k + \varphi(\log d_{jk}) + \varepsilon_{jk} \tag{4.1}$$

[14] The 95% confidence interval for observation i is calculated as 1.96 times the robust standard error of the intercept in the local kernel regression centered on observation i.

[15] Virtually identical figures obtain if we use only non-migrant households.

where B_{jk} is a vector of controls for household j in location k and $\varphi(.)$ is an arbitrary smooth function. The composition of the household is captured by the number of household members, a female head dummy, and the shares of women, young children, youth, and elderly members in the household.[16] We also include the age and age squared of the head and we controls for price differentials to the extent allowed by the data. Prices controls include the ward median rice price, the ward median wage, and regional dummies.

The estimated function $\varphi(\log d_j)$ is depicted in Figure 2. Results confirm the existence of a strong negative relationship between isolation and consumption expenditures. This finding could be due to sorting on ability across locations, as discussed in the conceptual section, or it could be because isolation from markets offers fewer income earning opportunities. Since consumption is lower in isolated households, this could explain lower reported satisfaction level. It is therefore important that we control for expenditures when measuring the relationship between isolation and subjective welfare.

4.2. Multivariate analysis

We now turn to a multivariate analysis. We begin by estimating an empirical equivalent of equation (2.3):

$$W_{jk}^h = f(\alpha_0^h + \alpha_1^h \log X_{jk} + \alpha_2^h \log d_{jk} + \alpha_3^h \log P_k + \alpha_4^h D_k + \alpha_5^h \log p_k + \alpha_6^h B_{jk}) \qquad (4.2)$$

where W_{jk}^h denotes the satisfaction rankings discussed earlier and $f(.)$ is an ordered probit density function. Our first isolation variable is distance to markets d_{jk}. Urban population within two-hour travel time from the ward, P_k, and population density in the district D_k are included as additional measures of isolation: households living in sparsely populated districts on

[16] Adult males are the omitted category.

average live further away from each other. In equation (4.2), coefficients α_2^h, α_3^h and α_4^h proxy for the combined effect of amenities, product variety, and local public goods.

The log of consumption expenditure $\log X_{jk}$ is included as regressor.[17] As in (4.1), controls B_j are included to correct for differences in household size and composition. Household size is expected to reduce income and consumption adequacy because the same level of expenditures should yield less satisfaction if the household is larger. The age and gender composition of the household may also affect how much satisfaction is derived from a given level of consumption expenditures.[18] Age is included to allow for life cycle effects: we expect young people to be less satisfied with life in general if their expectations are inflated by the prospect of economic growth. We expect the female head dummy to have a negative coefficient because many female headed households result from divorce and separation. The log of the value of household assets is included as additional regressor to capture permanent income effects hidden by a transitory rise or fall in expenditures. Assets may also affect subjective well-being directly through the sense of security they provide (e.g. Deaton 1991).

The median rice price and wage rate in the ward are included as price controls. We also include a district-specific housing price premium. This is estimated by regressing the (log of the) monthly rental price on district dummies, controlling for a variety of house characteristics such as square footage, number and type of rooms, quality of materials, and in-house amenities. District dummies are thought of as capturing locational attributes such as access to public amenities and the like. We therefore expect subjective welfare to increase with the locational

[17] We experimented with a more flexible cubic functional form in log consumption but got virtually identical results regarding isolation. This is hardly surprising given that, as shown in Figure 2, the relationship between $\log X_{jk}$ and $\log d_{jk}$ is basically linear.

[18] In particular, we note that female members typically produce services (such as home care, knitting, and sewing) which are consumed by the household (Fafchamps & Quisumbing 2003). Because it is extremely difficult to impute a value on such services, they are omitted from consumption expenditures. Adult males, in contrast, typically focus on self-employment or wage work. The monetary income they bring is properly measured as part of consumption expenditures. For this reason, we expect the share of female members in the household to raise subjective well-being once we control for consumption expenditures. These effects are captured by the share of various age/sex groups in the regression.

premium.

Multivariate regression results are presented in Table 4 using ordered probit.[19] Robust standard errors are reported throughout, clustered by ward to allow for correlation in errors for households residing in the same location. All regressions show a negative effect of distance to markets d_{jk} on subjective satisfaction. The effect is strong and significant in five of the six regression, the exception being the total income regression. Our second measure of isolation P_k is positive and strongly significant in all six regressions. If we omit P_k from the income regression, the coefficient of d_{jk} is significant. Taken together, these results imply that, after controlling for consumption expenditures and household composition, subjective satisfaction is higher in households located close to markets and in or nearby large urban centers. It is not just distance to local markets that matters, but also the size of the urban population in nearby towns. Our third measure of isolation, population density D_k, is positive and significant at the 10% level or better in four of the six regressions, further confirming the relationship between subjective welfare and isolation. Population density, however, has a negative and significant effect on housing adequacy. This is probably due to a price effect as population concentration raises rents and house prices.

Taken together, these results indicate that subjective welfare is negatively correlated with isolation even after factoring out the effect of lower consumption expenditures. The regression results also shed some indirect light on the nature of isolation-welfare relationship. Normalized distance coefficients α_2^h/α_1^h are reported at the bottom of Table 4 together with their t-value. Results indicate that the relative magnitude of the distance coefficient is largest for health care and, to a lesser extent, for schooling and housing. This is probably because households living in

[19] Given that our estimator is probit, we could in principle achieve a gain in efficiency by estimating all six regressions as a seemingly unrelated system of regressions, thereby allowing errors to be correlated across equations. Given that dependent variables are categorical, this would require six levels of numerical integration – a feat of computer programming that is not justified by the anticipated efficiency gain.

isolated wards find it difficult to obtain health care in case of medical emergencies. This suggests that access to public services may be a large component of the cost of isolation.

We also find that distance coefficients are larger for questions relating to satisfaction with consumption than for the income question itself. This suggests that answers to the income adequacy questions do not fully capture the non-monetary costs of isolation, such as lower product variety and access to public services. If this interpretation is correct, it follows that most welfare costs of isolation are non-monetary. We revisit this issue in the next section.

Turning to other regressors, most controls have the anticipated sign. We find a positive and significant coefficient of consumption expenditures in the regressions for all income and consumption adequacy questions. Household assets have a significant positive coefficient in all six regressions while household size has a significant negative coefficient in most of them. These results are consistent with the utility model. In most cases, household size and consumption expenditure – both of which are in logs – have roughly the same coefficient, except with a different sign. Results would thus not change much if we simply divided consumption by the number of household members instead of entering both regressors independently.

The locational housing price premium has the anticipated positive sign and is significant in all regressions. We also note that the distance coefficient is larger when the housing price premium is omitted from the regression, suggesting that some of the effects of isolation are captured by the housing price variable. Other village-level prices have a negative and significant effect on satisfaction from food consumption, but in other regressions the price variables are mostly non-significant. We also find strong regional differences. With the exception of health care, households located in the Mountain and Hills zone tend to report lower levels of satisfaction. This is again consistent with other isolation results: the steeper the terrain, the less likely travel is to take place on motorized vehicles, and the more arduous travelling to the market becomes.

4.3. Possible self-selection bias

The utility approach does not depend on whether people are mobile or not – and hence is not affected by selection across locations according to ability. But there may be unobserved individual characteristics other than ability that influence subjective utility and are correlated with distance. For instance, it is conceivable that there exist 'grumpy' people who tend to be less intrinsically happy. As they are less sociable, they self-select into remote locations. This is a potential source of bias.

Since the bias arises from self-selection, mobility is at the heart of the econometric problem: among people who cannot move, there should be no systematic relationship between 'grumpiness' and isolation as long as grumpiness is a randomly distributed human trait. In our data, 80% of the surveyed heads of household reside in their birth ward. This suggests a strategy for dealing with potential self-selection bias. We estimate equation (4.2) using only non-migrant households and correct for self-selection into migrant status as follows:

$$\begin{aligned} W^h_{jk} &= \alpha^h_0 + \alpha^h_1 \log X_{jk} + \alpha^h_2 \log d_{jk} + ... + u^h_{jk} \text{ if } M_j{=}0 \qquad (4.3) \\ M_j &= 1 \text{ if } M^*_j = \rho Z_j + v_j \geq 0 \\ &= 0 \text{ if } M^*_j = \rho Z_j + v_j < 0 \end{aligned}$$

In the country of study, male adults migrate early in life (Seddon, Adhikari & Gurung 1999). Migrant household heads are those who were surveyed in a ward other than their birth ward. The regressors Z_j are variables affecting the decision to leave one's birth place. They include predetermined individual characteristics such as education of the head and parental education. Inherited land is included as well because it is tied to location specific knowledge that would be lost if the household were to move. Date of birth is included to reflect changes in migration opportunities over time. Ethnicity dummies are included in case certain groups have better networks with migrant populations elsewhere (e.g. Seddon, Adhikari & Gurung 1999, Munshi 2003). In the following sub-section we also include ethnicity dummies and education of the household head as additional controls in the subjective welfare regressions. Inherited land and education and occupation of the father thus serve to identify the selection equation. They are reasonable instruments for our purpose since they are likely to affect the migration decision but unlikely to affect 'grumpiness' per se.

Model (4.3) is estimated using the bivariate probit selection estimator of Heckman. To this effect, we recode answers to the satisfaction question into two categories only – less than adequate, and adequate or more than adequate. This entails a loss of information but since the number of observations in the 'more than adequate' category is very small, the loss of information is minimal.

The selection regressions are presented in Appendix.[20] They show that better educated heads of household are significantly less likely to have remained in their birth ward. This is consistent with empirical evidence showing that returns to education are highest in non-farm activities (e.g. Yang 1997, Fafchamps & Quisumbing 2003). We also find that migrating out is more likely if the head's father is better educated, possibly reflecting an interest in off-farm work from an early age. In contrast, households inheriting a lot of land from their parents are less likely to have migrated out of their birth ward. These two variables are strongly significant, confirming that the selection equation is identified. Several ethnic dummies are also significant, with those belonging to the Brahmin caste and to the Magar and Tharu tribes more likely to migrate.

Regression results for W^h_{jk} are presented in Table 5. Using a likelihood ratio test, the absence

[20] Since the selection and adequacy regressions are estimated jointly, there is one selection regression per adequacy regression. Estimated coefficients are very similar across regressions.

of correlation between the errors in the selection and satisfaction regressions is only rejected for the food regressions – but it would be rejected at p-values of 20% or less, except for health care. A selection correction is thus appropriate. As is clear from Table 5, our main results regarding isolation are basically unchanged: distance is significantly negative in all regressions except total income. The consumption expenditure variable remains positive and significant except for health care, where it is now non-significant. Other qualitative results survive as well. Thus, although a selection correction may be appropriate in this case, self-selection does not appear to be responsible for our findings regarding the welfare cost of isolation.

4.4. Robustness checks

The theoretical model presented in Section 2 suggested that isolation may affect welfare through various channels, such as prices p, access to public goods A, and variety of consumer goods and services N. Unfortunately we only have partial information about these channels. In Tables 4 and 5 we have already made use of the limited price information available. The results have shown that, as predicted by theory, subjective satisfaction with food consumption is lower when the local price of rice is higher. In contrast, our housing price index has a positive – and often significant – coefficient in all regressions, suggesting that the variable proxies for various locational advantages. This finding is consistent with the work on Jacoby (2000) who found land prices to fall with isolation in Nepal.

We now include additional variables that proxy for N and A. As proxy for A, we include the (log of the) distances from the household to the nearest school and health facility. If the relationship between isolation on subjective welfare is driven by differences in access to schools and health care, introducing these variables in the regression should result in a zero coefficient of isolation variables $d_{jk}, \log P_k$ and D_k – especially in the schooling and health care regressions.

As proxy variables for N, we use the three indices of variety N^J discussed in the data section. Although imperfect, these measures give an idea of the number of distinct categories of products and services available to ward residents.

To minimize the risk of omitted variable bias, we also add a number of regressors thought to affect subjective welfare. We begin by adding the education of the household head. Education has been shown to influence responses to subjective welfare questions (Diener, Suh, Lucas & Smith 1999). Unemployment and illness are included for similar reasons. The Gini coefficient of consumption per capita in the ward is included to capture possible aversion to inequality. Rainfall in the year preceding the survey and the ward-specific standard deviation of rainfall in the year of the survey are included to capture possible effects of climate on residents' mood. To capture possible effects of the Maoist insurrection on people's expectations, we include our insurgency dummies.[21] To control for social status we include ethnic dummies. Finally we include a dummy variable taking value 1 if the household hired permanent or casual workers in the year of the survey. The rationale for doing so is that households employing other people may feel they enjoy a higher status, and this may affect their response to adequacy questions.[22] Results with the additional regressors are summarized in Table 6. A selection correction is conducted in the same way as before but not shown here to save space.[23]

Additional controls for isolation fall short of expectations. Distance to the nearest school and health facility are never significant, suggesting that differences in physical distance to these facilities do not account for the relationship between isolation and subjective welfare. Other

[21] Admittedly, the information at our disposal measures the incidence of the insurgency four years after the survey. However, it is likely that over time the insurgency got strongest in the areas in which its action was already perceived in early 1996 when actions started. Insurgency dummies can thus be seen as an effort to capture the insurrective mood of the population in 1996.

[22] Adding this variable may also clarify the effect of the wage variable since it is likely to differ if the household is a buyer or seller of labor.

[23] After including of the additional regressors, the null hypothesis of no correlation between the error terms in the selection and satisfaction regressions can only be rejected for the food regression.

dimensions of local public service provision probably matter more, such as the quality of the school or health facility and the availability of drugs and teachers, for which we do not have data.

Indices N^J of product variety are significant in a number of regressions – mostly the index of non-food consumption. We take this as evidence that product variety is valued by Nepalese households. However, the effect of the inclusion of these variables on the distance coefficients is minimal, suggesting that the N^J indices are far from accounting for the effect of isolation.

Regarding the isolation variables themselves, our main results are basically unchanged: distance remains negative and significant in all regressions except total income while urban population remains positive and significant in all regressions. Population density remains significant in two regressions. Other controls need not be discussed in detail since their inclusion is purely to eliminate possible sources of omitted variable bias.[24]

These results demonstrate that the relationship between isolation and subjective adequacy survives the elimination of many potential sources of omitted variable bias. But they also indicate that we have not been able to identify the precise channel through which geographical isolation and subjective welfare are related.

As another robustness check, we investigate whether our results may be affected by endogenous placement within wards. Tables 5 and 6 control for self-selection *across* wards. The reader may nevertheless worry about the possible endogeneity of household placement *within* the ward

[24] Education of the household head is positive and significant and unemployment is negative and significant in all regressions. Illness is negative in all regression, significantly so in five. These results are in line with experimental evidence (e.g. Frey & Stutzer 2002, Diener & Biswas-Diener 2000). We find that more rain tends to make people less satisfied (significant in three regressions), perhaps because rains damage roads and isolate wards further. Ethnicity variables are significant in a few regressions, usually suggesting that members of some of the tribal groups are more easily dissatisfied, perhaps because of political grievances. The labor hiring variable is marginally significant in two regressions. The Gini coefficient is significant in one regression but with the wrong sign. Maoist insurgency coefficients, when significant, usually have the wrong sign, with more affected regions appearing to be more satisfied with their income and consumption than inhabitants of least affected areas. Whatever the explanation for these results, they demonstrate that inequality and the Maoist insurgency are not what accounts for the negative relationship between income and consumption adequacy and isolation.

– e.g., that grumpy people live at the outskirts of the village. To investigate this possibility, we reestimate Table 6 replacing individual distance d_{jk} with the ward average $\overline{d}_k$.[25] To save space, we show in Table 7 the regression results for the distance coefficients only. Distance is even more significant, indicating that our earlier results are not driven by endogenous placement within the ward.

We also estimate the regression including both household-specific distance to market and average distance in the ward. Multicollinearity between the two measures gets in the way of precise estimation. The results, not presented here to save space, show that ward average distance is negative and significant at the 10% level or better in four of the six regressions. What matters most appears to be isolation of the ward itself, not relative isolation of individuals within the ward. This is further evidence that endogenous placement within the ward is unlikely to account for our results.

5. Magnitude

We have found a robust and significant relationship between isolation and subjective welfare. But is the magnitude of the relationship large enough to warrant further consideration? To quantify it, we draw upon the formula derived in Section 2 for estimating the equivalent variation c_{km} of reducing travel time from, say, d_k to d_m:

$$c_{km} = 1 - e^{\frac{\alpha_2}{\alpha_1}(\log d_k - \log d_m)} \tag{5.1}$$

This formula provides a useful yardstick for quantifying the magnitude of the relationship between d_k and subjective welfare.

We compute formula (5.1) replacing α_1 and α_2 by the coefficients of distance and con-

[25] The ward mean is computed excluding the household itself, so as to avoid spurious correlation.

sumption expenditures. This provides an intuitive way of quantifying the relationship between distance and welfare: if the relationship between d_k and subjective welfare could be interpreted as causal, c_{km} would measure the subjective cost of isolation in monetary terms.[26] Each of our six regressions yields separate α_1 and α_2 estimates and hence a different c_{km}. Differences among these c_{km}'s gives an idea of the relative magnitude of the welfare cost of isolation on different components of utility.

The coefficient of income is not significantly different from 0 in the health care regression. This is consistent with rationing in health care, as would be the case if health services are subsidized by government. As argued in the conceptual section, equation (5.1) is no longer valid if there is quantity rationing but we may be able to bracket c_{km} if we are willing to assume that $\lambda_r \approx \frac{\alpha_2^r}{\alpha_1^u}$. In this case, we can use income coefficients estimated for unrationed consumption goods to normalize the distance coefficient in the rationed regression. This means calculating (5.1) with four different α_1 and reporting the range of values found. Given that the estimated α_1^u are broadly similar across categories, we expect respondents to have answered all the adequacy questions in a comparable way.

If preferences are (approximately) homothetic, we can also compute the combined effect of isolation using $V = \sum_{h=1}^{H} \omega_h V^h$ where ω_h is the consumption share of subset h. As before, we can write $V^h = b' + \log X - \lambda_h \log d$ where $\lambda_h = \alpha_2^h/\alpha_1^h$. We obtain the combined welfare cost of isolation by solving $V_k = V_m$ which yields:

$$\begin{aligned} \sum_{h=1}^{H} \omega_h(b' + \log X - \lambda_h \log d_k) &= \sum_{h=1}^{H} \omega_h(b' + \log X(1 - c_{km}) - \lambda_h \log d_m) \\ c_{km} &= 1 - \exp\left(\sum_{h=1}^{H} \omega_h \frac{\alpha_2^h}{\alpha_1^h} (\log d_k - \log d_m)\right) \qquad (5.2) \end{aligned}$$

[26] Since, for a given $\log d_k - \log d_m$, c_{km} is a non-linear combination of parameter estimates, a confidence interval can be computed as well.

where we have used the fact that consumption shares ω_h sum to one. Equation (5.2) says that the combined welfare cost of isolation is a weighted combination of effects on consumption subsets. Because of suspected rationing in health care, we again use the bracketing method for health care, yielding a range of possible values for c_{km}.

As is clear from (5.1) and (5.2), computing the welfare cost of isolation ultimately involves dividing the distance coefficient α_2^h by the consumption expenditure coefficient α_1^h. Estimating the magnitude of the relationship between isolation and welfare thus requires obtaining a consistent estimate of the consumption expenditure coefficient α_1^h. So far we have focused primarily on obtaining a consistent estimate of α_2^h because our objective was to test the existence of a subjective welfare cost to isolation. Now we also need to worry about α_1^h. Because of attenuation bias, measurement error in consumption expenditures leads to an underestimation of α_1^h and thus an overestimation of c_{km}.

To correct for measurement error, we need to instrument consumption expenditures. In selecting instruments, we must avoid variables that may be correlated with the error term in the adequacy regressions. For instance, it is conceivable that individuals with a grumpy disposition earn less than cheerful individuals, and hence consume less. For this reason, instruments must not include variables possibly correlated with grumpiness, such as household size or current assets.[27] To this effect, we only use variables that can reasonably be regarded as pre-determined from the individual's perspective – such as parental background, age, education, and ethnicity. Of those, only parental background can reasonably be omitted from the consumption adequacy regression. Occupational choice has a strong effect on income – especially farm versus non-farm – but it is possibly endogenous, so we cannot control for it directly. However, we can control for it indirectly as follows. Regression results reported in Table A2 make us suspect that

[27]People suffering from depression, for instance, often antagonize those around them and probably earn less as a result.

children born to educated parents involved in non-farm work are less likely to work in agriculture. Parental background can thus instrument for occupational choice. Furthermore, the income of agricultural households depends more strongly on the level and variation of rainfall than that of non-agricultural households. Following the same approach as Fafchamps, Udry & Czukas (1998), the effect of local weather conditions on agricultural income can thus be instrumented by interacting the mean and standard deviation of district rainfall with the education and occupation of the head's father.

The instrumenting equation is shown in Table 8. Our instruments are jointly significant and the F-statistic is only marginally below 10, which is considered sufficient to avoid a weak instrument problem. We also conduct an overidentification test, temporarily ignoring the selection issue. In all cases, we cannot reject the null hypothesis that instruments can be excluded from the adequacy regressions: p-values are 34% and above.

The model presented in Table 7 is reestimated with all regressors and controls, adding the residuals from the instrumenting equation in the adequacy regression. This approach is similar to that developed by Smith & Blundell (1986) for tobit and by Rivers & Vuong (1988) for logit. By analogy, it should also work here since the Heckman selection model is also based on the normal distribution. This approach yields a test of endogeneity as a by-product. Results are summarized in Table 9. Residuals from the instrumenting equation are significant in all regressions except health care, suggesting that endogeneity is indeed a problem. The main change compared to Table 7 is the massive increase in the consumption coefficients. This is a typical outcome when correcting for measurement error. Distance coefficients remain by and large unchanged.

Coefficient estimates from Table 9 are used to compute (5.1) and (5.2). The first calculations we report are the estimates and 95% confidence interval of the compensating variation corresponding to a reduction in distance from the 75th to the 25th percentile. This corresponds to a fall of just below two hours of travel time to the nearest market. For food consumption, clothing, housing, schooling, and total income we use formula (5.1). For health care we bracket c_{km} using $\lambda_r \approx \alpha_2^r/\alpha_1^u$ as explained above. We also compute the combined welfare cost of isolation using (5.2).[28]

Results suggest that the magnitude of the relationship between isolation and subjective welfare is quite large: the reduction in travel time from the 75th to the 25th percentile is associated with a compensative variation equivalent to 13.7% to 15.7% of current consumption expenditures. In terms of consumption subsets, the welfare gain would be highest for health care, housing, and schooling. But it would be smaller – and non-significant at the 95% level – when we use the answer to the total income adequacy question. As explained earlier, this is probably because respondents mentally distinguish between financial and access issues.

Two other sets are calculations are reported in the last two columns of Table 10. The first evaluates the equivalent variation of reducing travel time to markets from the mean of 2 hours and 10 minutes to the minimum recorded travel time, which is 1 minute. The second set evaluates the equivalent variation of the same reduction in travel time (2 hours and 9 minutes) for a household at the 90th percentile in terms of travel time. For such a household, travel time to markets is 5 hours and 20 minutes. The second set of estimates therefore represents the welfare gain of reducing travel time to markets from 5.34 to 3.18 decimal hours.

These calculations illustrate the non-linear relationship between isolation and subjective consumption adequacy. Completely eliminating geographical isolation generates a large subjective welfare gain: a household would be willing to forego 34-35% of its income to relocate from the

[28] In the surveyed population, average expenditure shares are as follow: food 66.3%; clothing 8.1%; housing 12.2%; schooling 2.8%; health 3.4%; other 7.2%. Adequacy questions thus cover items representing 92.8% of total consumption. Since we do not have an adequacy question for other goods, we ignore them in the calculation and renormalize shares to sum to 1. This is equivalent to assuming average subjective adequacy for other goods.

mean distance of 2.18 hours to the immediate vicinity of markets. The implied welfare gain from reducing isolation is much smaller when reducing travel time for household at the 90th percentile travel time. A household located more than 5 hours away from markets would only forego approximately 4.5% of its consumption to reduce travel time by the same 2 hours and 9 minutes. This is because, as illustrated in Figure 1, answers to adequacy questions are linear in $\log d_k$, not in travel time itself. As a result, the welfare gain from reducing travel time falls rapidly with distance. What matters the most is immediate vicinity to markets. A reduction in travel time from 15 to 5 minutes is as valuable as a reduction from 3 hours to 1 hour.

6. Discussion

We have shown that there is a significant and large relationship between geographical isolation and answers to consumption adequacy questions among Nepalese households. But we have been less successful at identifying the reason for this relationship.

Economic theory suggests several channels through which isolation may affect utility – such as income, the price and variety of consumption goods, and access to amenities and public goods. We did the best we could to control for these effects with the data at hand – and enjoyed some success in doing so. Income, proxied by total consumption expenditures, falls strongly with isolation and has a large and significant effect on reported adequacy of consumption – except for health care, where we suspect that for many households consumption is constrained by limited availability. The rice price has the predicted negative sign in the food adequacy regression, and the Herfindhal index for non-food consumption is positive in all regressions and significant in four. But other variables, such as distance to the nearest school or health facility, are largely non-significant – including, surprisingly, in the schooling and health care regressions. Better data is needed to identify the precise channels through which isolation affects subjective welfare.

So far we have proceeded as if answers to consumption adequacy questions are good proxies for utility. What if they are not? The empirical literature in psychology and economics concludes that people answer subjective well-being questions with a reference point in mind. This reference point may change over time and according to surrounding circumstances (e.g. Frey & Stutzer 2002, Kahneman, Diener & Schwarz 1999).

One important source of concern is what the psychology literature has called 'habituation', that is, the fact that human beings tend to judge their well-being by reference to past consumption (e.g. Kahneman, Diener & Schwarz 1999, Blanchflower & Oswald 2004). A rise in consumption initially increases subjective satisfaction but, over time, the new consumption level becomes the reference point. This idea has been applied by Pradhan & Ravallion (2000) to questions about consumption adequacy in Nepal. In the context of our modelling framework, it can formalized as follows:

$$V_{ht} = \log \frac{c_{ht}}{c_{hr}}$$

where c_{hr} is the reference point. If c_{hr} adjusts fully and instantaneously to current consumption, we have $c_{hr} = c_{ht}$ and subjective satisfaction remains constant irrespective of consumption. If c_{hr} adjusts only partially or with a lag, subjective satisfaction responds, but only partially, to consumption. For instance, if $c_{hr} = c_{ht}^{\phi}$ with $\phi < 1$, we have

$$\begin{aligned} V_{ht} &= \log \frac{c_{ht}}{c_{ht}^{\phi}} = (1-\phi)\log c_{ht} \\ &= (1-\phi)b' + (1-\phi)\log X - (1-\phi)\lambda_h \log d \end{aligned}$$

Habituation leads to a shrinkage of all coefficients: the stronger habituation is, the larger ϕ, and the smaller the coefficient of consumption X and that of distance d. The fact that we find a large coefficient on $\log d$ indicates that, however strong habituation is, it is not sufficient to eliminate the relationship between subjective welfare and geographic isolation.

The reference point c_{hr} may also vary with the consumption level of a reference group. According to psychologists, people derive satisfaction from their achievements which they judge in comparison to that of their peers, that is, of individuals who started life in similar conditions (e.g. Kahneman, Diener & Schwarz 1999, Layard 2002, Luttmer 2005). In the context of Nepal, this typically means people born in the same village. Fafchamps & Shilpi (2008) examine this issue in detail using the same data. They estimate a model in which c_{hr} is approximated by the average or median consumption levels of other households in the same ward. They show that answers to subjective consumption adequacy questions depend on consumption relative to others in the ward of residence and, for migrants, in the birth district. This result obtains even though Fafchamps and Shilpi also control for distance to the nearest market.

The relationship between isolation and subjective welfare therefore does not appear to depend on whether one controls for relative consumption or not. This is not surprising. We have seen that consumption expenditures fall with isolation. This means that geographically isolated households are surrounded on average by households with a low consumption level – and can thus be expected to have a lower reference point. This should raise the subjective welfare of isolated households – and hence cannot account for the negative relationship between isolation and subjective consumption adequacy. From this we conclude that habituation and reference point considerations are very unlikely to account for our findings: if anything, they should bias the coefficient of d_k upwards, i.e., towards zero.

7. Conclusion

Using 1995/96 household survey data from Nepal, we have estimated the relationship between geographical isolation and subjective welfare. This estimation rests on the assumption that responses to questions about income and consumption adequacy capture utility rankings. Nepal is a perfect country to study isolation because road construction is recent and much of the country remained inaccessible by road at the time of the survey.

We find the relationship between isolation and subjective welfare is significant and large in magnitude. Comparing the welfare cost of isolation across categories of consumption goods indicates that respondents factor the utility gain from product variety in their reported adequacy of consumption. Geographical isolation is associated with lower subjective consumption adequacy also for schooling and health care. In fact, for health care, total expenditures are not a significant determinant of access to health care, but isolation is. These findings suggest that welfare assessments based on geographical poverty maps (e.g. Ferreira, Lanjouw & Neri 2003, Alderman, Babita, Demombynes, Makhatha & Ozler 2002, Mistiaen, Ozler, Razafimanantena & Razafindravonona 2002) may underestimate the subjective welfare cost of isolation since these maps typically focus solely on monetary income and consumption.

Our results imply that, given time and opportunity, many rural dwellers may prefer to move out of isolated rural communities. This prediction is indeed happening. According to the 1991 population census, the capital city Katmandu counted less than half a million people. The population census conducted a decade later indicates that it now has about three times as many inhabitants. The number and population of small towns has also increased dramatically. The evidence provided in this paper suggests that rural dwellers are attracted by the amenities and lifestyle that urban centers provide – proximity to markets, variety of goods and services, better access to schools and health care. This phenomenon may explain why countries that have seen little growth and thus little employment 'pull' from cities – such as many parts of Sub-Saharan Africa – have nevertheless experienced massive urbanization.

References

Alderman, Harold, Miriam Babita, Gabriel Demombynes, Nthabiseng Makhatha & Berk Ozler. 2002. "How Long Can You Go? Combining Census and Survey Data for Mapping Poverty in South Africa." *Journal of African Economies* 11(2):169–200.

Blanchflower, David G. & Andrew J. Oswald. 2004. "Well-Being over Time in Britain and the USA." *Journal of Public Economics* 88(7-8).

Dahl, Gordon B. 2002. "Mobility and the Return to Education: Testing a Roy Model with Multiple Markets." *Econometrica* 70(6):2367–2420.

Deaton, Angus. 1991. "Saving and Liquidity Constraints." *Econometrica* 59(5):1221–1248.

Deaton, Angus & Christina H. Paxson. 1998. "Economies of Scale, Household Size, and the Demand for Food." *Journal of Political Economy* 106(5):897–930.

Diener, Ed, Eunkook M. Suh, Richard E. Lucas & Heidi L. Smith. 1999. "Subjective Well-Being: Three Decades of Progress." *Psychology Bulletin* 125(2):276–303.

Diener, Ed & Robert Biswas-Diener. 2000. "New Directions in Subjective Well-Being Research: The Cutting Edge." (mimeograph).

Elbers, Chris, Jean O. Lanjouw & Peter Lanjouw. 2003. "Micro-Level Estimation of Poverty and Inequality." *Econometrica* 71(1):355–64.

Fafchamps, Marcel, Christopher Udry & Katherine Czukas. 1998. "Drought and Saving in West Africa: Are Livestock a Buffer Stock?" *Journal of Development Economics* 55(2):273–305.

Fafchamps, Marcel & Agnes R. Quisumbing. 2003. "Social Roles, Human Capital, and the Intra-household Division of Labor: Evidence from Pakistan." *Oxford Economic Papers* 55(1):36–80.

Fafchamps, Marcel & Forhad Shilpi. 2003. "The Spatial Division of Labor in Nepal." *Journal of Development Studies* 39(6):23–66.

Fafchamps, Marcel & Forhad Shilpi. 2005. "Cities and Specialization: Evidence from South Asia." *Economic Journal* . (forthcoming).

Fafchamps, Marcel & Forhad Shilpi. 2008. "Subjective Welfare, Isolation, and Relative Consumption." *Journal of Development Economics* . (forthcoming).

Ferreira, Francisco H.G., Peter Lanjouw & Marcelo Neri. 2003. "A Robust Poverty Profile for Brazil Using Multiple Data Sources." *Revista Brasileira de Economia* 57(1):59–92.

Frey, Bruno S. & Alois Stutzer. 2002. "What Can Economists Learn from Happiness Research?" *Journal of Economic Literature* XL:402–435.

Harris, John & M. Todaro. 1970. "Migration, Unemployment and Development: A Two-Sector Analysis." *Amer. Econ. Rev.* 60:126–142.

Jacoby, Hanan G. 2000. "Access to Markets and the Benefits of Rural Roads." *Economic Journal* 110(465):713–37.

Jalan, Jyotsna & Martin Ravallion. 2002. "Geographic Poverty Traps? A Micro Model of Consumption Growth in Rural China." *Journal of Applied Econometrics* 17(4):329–46.

Kahneman, Daniel, Ed Diener & N. Schwarz. 1999. *Well-Being: The Foundations of Hedonic Psychology.* New York: Russell Sage Foundation.

Layard, Richard. 2002. "Rethinking Public Economics: Implications of Rivalry and Habit." (mimeograph).

Lewis, W. Arthur. 1954. "Economic Development with Unlimited Supplies of Labour." *The Manchester School* XXII(2):139–191.

Luttmer, Erzo F.P. 2005. "Neighbors as Negatives: Relative Earnings and Well-Being." *Quarterly Journal of Economics* 120(3):963–1002.

Mistiaen, Johan A., Berk Ozler, Tiaray Razafimanantena & Jean Razafindravonona. 2002. Putting Welfare on the Map in Madagascar. Technical report African Region Working Paper No. 34, The World Bank Washington DC: .

Munshi, Kaivan. 2003. "Networks in the Modern Economy: Mexican Migrants in the US Labor Market." *Quarterly Journal of Economics* 118(2):549–99.

Murphy, Kevin M., Andrei Shleifer & Robert W. Vishny. 1989. "Industrialization and the Big Push." *J. Polit. Econ.* 97(5):1003–26.

Pradhan, Menno & Martin Ravallion. 2000. "Measuring Poverty Using Qualitative Perceptions of Consumption Adequacy." *Review of Economics and Statistics* 82(3):462–71.

Ravallion, Martin & Jyotsna Jalan. 1996. "Growth Divergence Due to Spatial Externalities." *Economic Letters* 53(2):227–32.

Ravallion, Martin & Jyotsna Jalan. 1999. "China's Lagging Areas." *Americal Economic Review* 89(2):301–05.

Ravallion, Martin & Quintin Wodon. 1999. "Poor Areas or Only Poor People?" *Journal of Regional Science* 39(4):689–711.

Rivers, D. & Q.H. Vuong. 1988. "Limited Information Estimators and Exogeneity Tests for Simultaneous Probit Models." *Journal of Econometrics* 39:347–66.

Roback, Jennifer. 1982. "Wages, Rents, and the Quality of Life." *Journal of Political Economy* 90(6):1257–78.

Rosen, Sherwin. 1979. Wages-based Indexes of Urban Quality of Life. In *Current Issues in Urban Economics*. Baltimore: Peter Mieszkowski and Mahlon Straszheim (eds.), Johns Hopkins University Press.

Roy, Andrew D. 1951. "Some Thoughts on the Distribution of Earnings." *Oxford Economic Papers* 3:135–46.

Seddon, David, Jagannath Adhikari & Ganesh Gurung. 1999. "Foreign Labour Migration and the Remittance Economy of Nepal." (mimeographed report to DfID).

Smith, R. & Richard Blundell. 1986. "An Exogeneity Test for a Simultaneous Equation Tobit Model with an Application to Labor Supply." *Econometrica* 54:679–685.

Yang, Dennis T. 1997. "Education and Off-Farm Work." *Economic Development and Cultural Change* 45 (3):613–632.

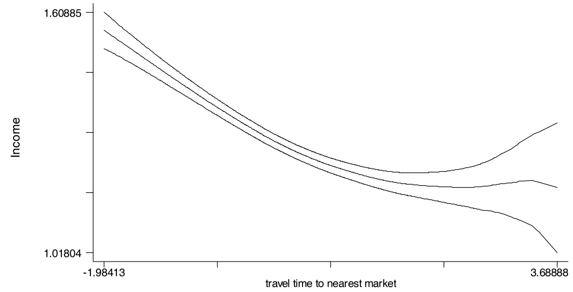

Figure 1a. Total income

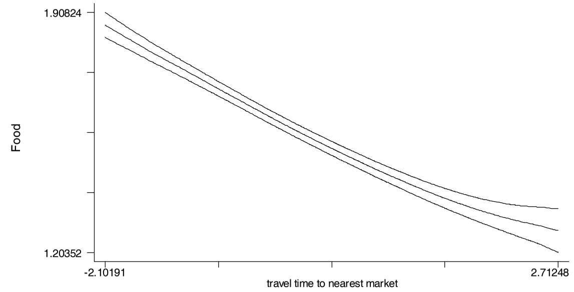

Figure 1b. Food consumption

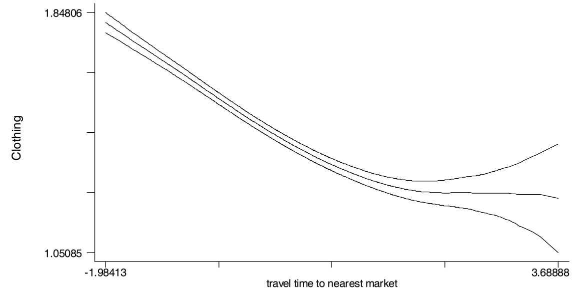

Figure 1c. Clothing consumption

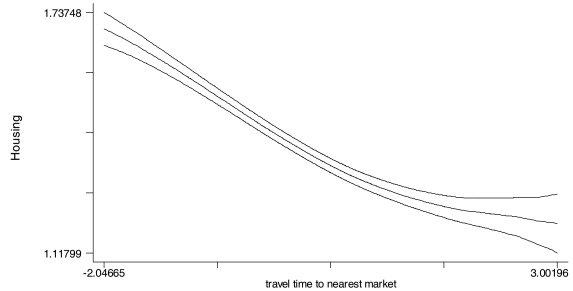

Figure 1d. Housing

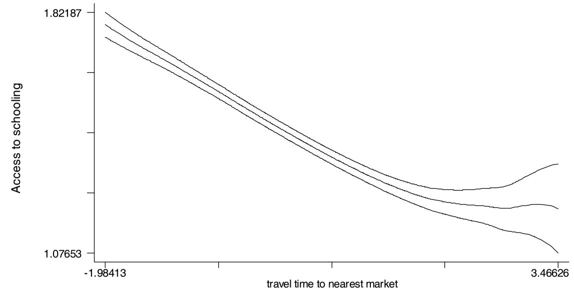

Figure 1e. Access to schooling

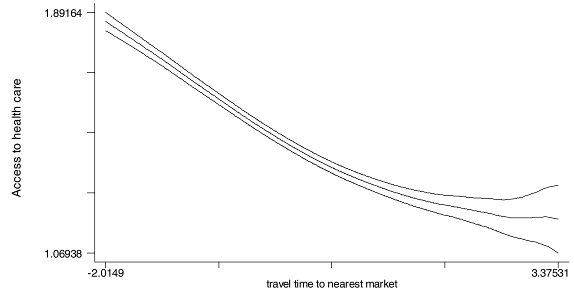

Figure 1f. Access to health care

Figure 1. Non-parametric analysis

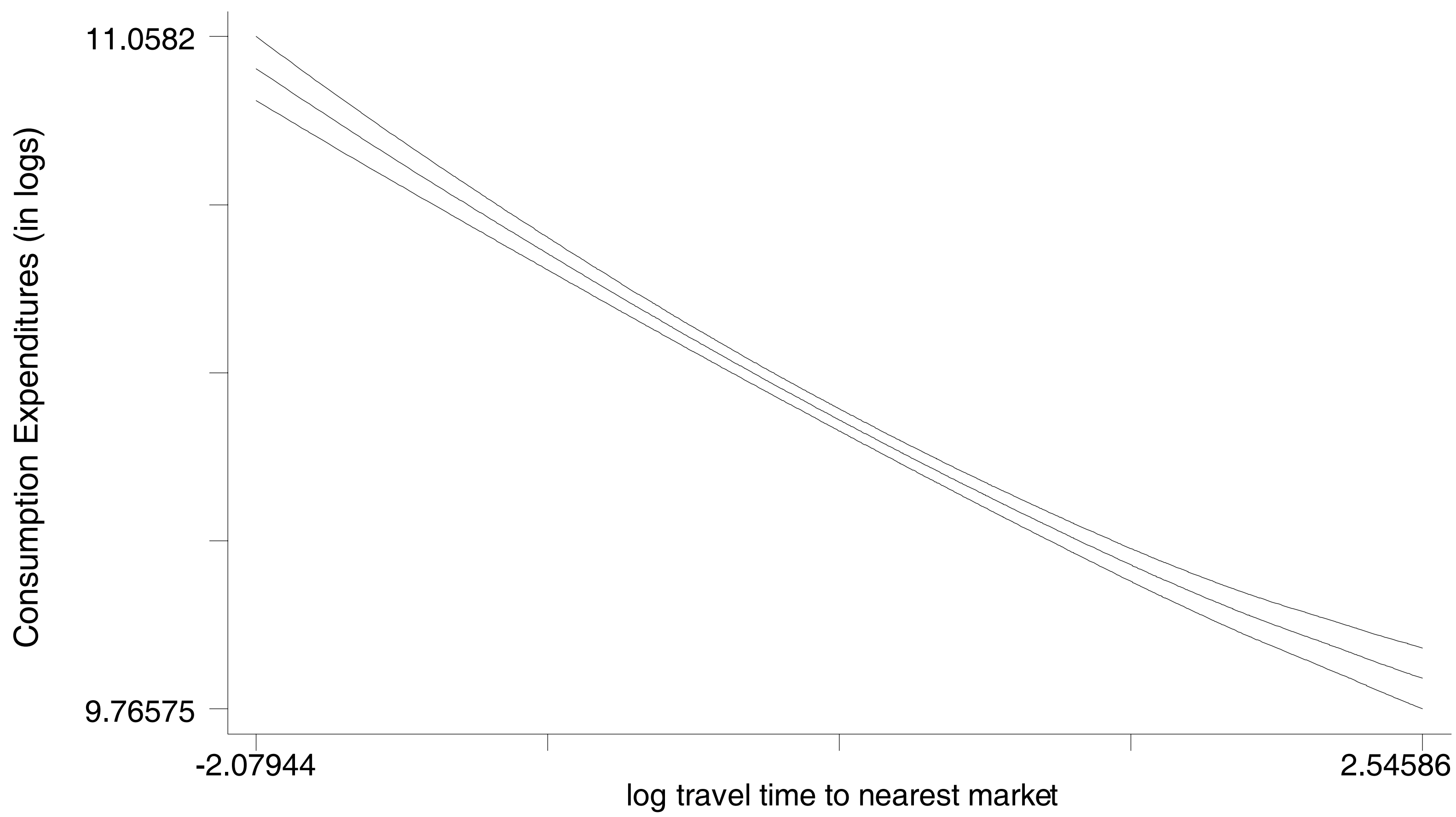

Figure 2. Consumption and isolation

Table 1. Answers to income and consumption adequacy questions

	Percent of responses: less than adequate	adequate	more than adequate
Total income	68.7%	30.6%	0.7%
Food consumption	46.6%	51.4%	2.0%
Clothing	52.7%	46.9%	0.3%
Housing	58.8%	41.0%	0.1%
Schooling	52.6%	47.1%	0.3%
Health care	52.0%	47.9%	0.1%
Number of observations	3317		

Table 2. Household characteristics

Isolation	**Unit**	**Mean**	**Median**	**St.dev.**	**Min.**	**Max.**
Travel time to nearest local market	Hours	2.18	1.06	3.36	0.01	40.00
Travel time to nearest school	Hours	0.37	0.25	0.88	0.00	25.00
Travel time to nearest health facility	Hours	1.08	0.50	1.78	0.00	30.00
Consumption						
Total annual consumption expenditures	US$	862	563	1015	29	19940
Total value of assets	US$	9910	2445	29854	0	714789
Household size and composition						
Number of household members	Number	5.6	5.0	2.8	1.0	29.0
Share of adult females in the household	Share	0.26	0.25	0.16	0.00	1.00
Share of children aged 6 and under	Share	0.15	0.13	0.16	0.00	0.67
Share of youths aged 7 to 20	Ahare	0.32	0.33	0.22	0.00	1.00
Share of members aged 65 and above	Share	0.04	0.00	0.13	0.00	1.00
% households with female head		13.6%				
Household characteristics						
Age of household head	Years	44.8	43.0	14.4	11.0	92.0
Years of schooling of household head	Years	3.6	0.0	5.0	0.0	17.0
Days lost to illness over preceding 12 month	Days	2.3	0	6.57	0	70
% hholds with 1 or more members unemployed		18.0%				
% hholds that hire permanent or casual labor		80.1%				
Parental background						
Inherited land	Hectares	0.81	0.36	1.69	0.00	32.05
Years of schooling of head's father	Years	0.9	0.0	2.5	0.0	16.0
% hholds in which head's father had non-farm job		17.0%				
Number of observations		3337				

Table 3. Ward variables

Isolation	Unit	Mean	Median	St.dev.	Min.	Max.
Urban population within 2 hours travel time	thousands	128.0	0.0	218.0	0.0	795.0
Population density in the district	per sqkm	383	185	483	2	1692
Prices						
House rental price	US$/year	553.9	411.8	474.17	15.4	2155
Median rice price in ward	US$/Kg	0.44	0.44	0.14	0.12	1.04
Median wage rate in ward	US$/day	0.75	0.48	1.01	0.00	12.35
Inequality						
Inequality in household consumption	Gini coef.	0.290	0.282	0.083	0.083	0.609
Inequality in per capita consumption	Gini coef.	0.257	0.246	0.082	0.091	0.509
Indices of consumption variety						
Herfindahl index for food products	Index	0.12	0.09	0.07	0.03	0.49
Herfindahl index for non-food products	Index	0.14	0.14	0.03	0.09	0.42
Herfindahl index for durable goods	Index	0.44	0.44	0.16	0.15	0.85
Rainfall						
Average rainfall in ward	mm	1702	1459	612	1039	3431
Standard deviation of rainfall in ward	mm	411	366	197	176	903
Maoist Insurgency as reported in 2000						
In a most affected district		2.5%				
In an affected district		9.9%				
In a little affected district		20.4%				
Not affected		67.2%				
Number of wards		274				

Table 4. Regressing income and consumption adequacy on isolation and household characteristics

	Subjective adequacy of:					
Isolation	**Food**	**Clothing**	**Housing**	**Schooling**	**Health care**	**Income**
Travel time to nearest market (log)	-0.188	-0.124	-0.153	-0.205	-0.254	-0.048
	(4.58)**	(2.93)**	(3.28)**	(4.27)**	(5.54)**	-1.16
Urban population within 2 hrs travel time	0.853	0.636	1.372	0.867	1.206	0.832
	(3.25)**	(2.52)*	(4.47)**	(3.07)**	(3.32)**	(3.41)**
Population Density (per sqkm)	0.001	0.000	0.000	0.000	0.000	0.000
	(3.83)**	(2.59)**	(2.23)*	(1.02)	(0.94)	(0.21)
Household consumption and controls						
Consumption expenditures (log)	0.591	0.517	0.233	0.339	0.164	0.352
	(7.30)**	(6.79)**	(3.18)**	(4.17)**	(2.17)*	(5.08)**
Value of assets (log)	0.161	0.127	0.105	0.109	0.098	0.121
	(6.28)**	(5.74)**	(4.84)**	(5.06)**	(5.30)**	(5.22)**
Household size (log)	-0.397	-0.345	-0.090	-0.336	-0.109	-0.198
	(4.30)**	(3.97)**	(1.05)	(3.56)**	(1.21)	(2.28)*
Share of adult females	0.101	0.186	0.030	0.270	0.110	0.032
	(0.38)	(0.68)	(0.11)	(0.81)	(0.40)	(0.11)
Share of children 6 and under	-0.308	0.086	-0.515	0.079	-0.379	-0.360
	(1.17)	(0.31)	(1.93)	(0.25)	(1.32)	(1.31)
Share of youths aged 7 to 20	-0.257	-0.259	-0.538	0.107	-0.268	-0.592
	(1.16)	(1.16)	(2.46)*	(0.41)	(1.20)	(2.52)*
Share of elderly 65 and above	-0.362	0.389	0.079	0.307	-0.188	0.144
	(1.32)	(1.28)	(0.27)	(0.77)	(0.72)	(0.50)
Age of household head	-0.017	-0.002	-0.015	-0.037	-0.017	0.004
	(1.53)	(0.21)	(1.29)	(2.86)**	(1.47)	(0.38)
Age of household head squared	0.000	0.000	0.000	0.000	0.000	0.000
	(1.55)	(0.20)	(1.26)	(2.73)**	(1.30)	(0.67)
Female head dummy	-0.053	0.003	0.037	-0.088	0.007	-0.068
	(0.58)	(0.04)	(0.40)	(0.84)	(0.07)	(0.73)
Prices						
House rental premium in district (predicted)	0.058	0.079	0.183	0.078	0.158	0.099
	(1.28)	(1.80)	(3.72)**	(1.42)	(2.96)**	(2.29)*
Median wage rate in ward (log)	-0.248	0.007	-0.065	-0.064	0.039	0.077
	(3.01)**	(0.10)	(0.75)	(0.68)	(0.46)	(1.05)
Median rice price in ward (log)	-0.281	-0.027	0.114	-0.101	-0.013	-0.167
	(1.69)	(0.18)	(0.76)	(0.55)	(0.08)	(1.25)
Regional and belt dummies			included but not shown			
Intercept	-5.798	-6.773	-4.172	-3.709	-3.453	-5.427
	(7.12)**	(8.83)**	(4.89)**	(3.89)**	(3.89)**	(7.39)**
Number of observations	3050	3048	3045	2451	3030	3041
Normalized coefficient of travel time to nearest market						
$\alpha 2/\alpha 1$	0.319	0.240	0.66	0.61	1.55	0.137
	(3.68)**	(2.59)**	(2.24)*	(2.85)**	(1.91)	(1.10)

Estimator is ordered probit. Absolute value of t statistics in parentheses: * significant at 5%; ** significant at 1%

Table 5. Adequacy regressions controlling for migration self-selection

	Subjective adequacy of:					
Isolation	**Food**	**Clothing**	**Housing**	**Schooling**	**Health care**	**Income**
Travel time to local markets (log)	-0.151	-0.065	-0.128	-0.214	-0.244	-0.051
	(3.42)**	(1.50)	(2.50)*	(3.98)**	(4.87)**	(1.10)
Urban population within 2 hrs travel time	1.188	0.920	1.777	1.301	1.789	0.913
	(4.12)**	(3.56)**	(5.23)**	(4.48)**	(4.95)**	(2.99)**
Population Density (per sqkm)	0.000	0.000	-0.001	0.000	0.000	0.000
	(1.31)	(0.36)	(2.94)**	(0.80)	(0.22)	(1.47)
Household consumption and controls						
Consumption expenditures (log)	0.501	0.393	0.136	0.189	0.065	0.257
	(5.76)**	(4.11)**	(1.63)	(1.95)	(0.70)	(2.94)**
Value of assets (log)	0.214	0.171	0.149	0.178	0.131	0.194
	(4.59)**	(4.67)**	(4.31)**	(4.67)**	(4.29)**	(4.68)**
Household size (log)	-0.349	-0.197	-0.003	-0.139	-0.011	-0.131
	(3.38)**	(1.93)	(0.03)	(1.19)	(0.10)	(1.26)
Share of adult females	-0.041	0.172	0.190	0.656	0.126	-0.188
	(0.12)	(0.58)	(0.64)	(1.63)	(0.39)	(0.57)
Share of children 6 and under	-0.403	-0.008	-0.561	0.156	-0.591	-0.772
	(1.29)	(0.03)	(1.77)	(0.40)	(1.82)	(2.41)*
Share of youths aged 7 to 20	-0.262	-0.166	-0.306	0.289	-0.271	-0.713
	(0.99)	(0.67)	(1.20)	(0.90)	(1.03)	(2.61)**
Share of elderly 65 and above	-0.347	0.283	-0.041	0.466	-0.177	-0.008
	(1.08)	(0.86)	(0.12)	(0.99)	(0.54)	(0.02)
Age of household head	-0.013	0.001	-0.017	-0.033	-0.019	-0.001
	(0.97)	(0.08)	(1.39)	(2.27)*	(1.41)	(0.09)
Age of household head squared	0.000	0.000	0.000	0.000	0.000	0.000
	(1.08)	(0.08)	(1.44)	(2.16)*	(1.30)	(0.10)
Female head dummy	0.073	0.075	0.004	-0.059	0.003	0.056
	(0.66)	(0.74)	(0.03)	(0.48)	(0.03)	(0.48)
Prices						
House rental premium in district (predicted)	0.034	0.038	0.151	0.066	0.114	0.093
	(0.75)	(0.88)	(2.93)**	(1.19)	(2.07)*	(2.03)*
Median wage rate in ward (log)	-0.195	-0.035	-0.081	-0.148	0.004	0.000
	(2.17)*	(0.44)	(0.83)	(1.50)	(0.04)	(0.00)
Median rice price in ward (log)	-0.247	-0.033	0.099	-0.273	-0.045	-0.243
	(1.44)	(0.22)	(0.59)	(1.31)	(0.24)	(1.70)
Regional and belt dummies			included but not shown			
Intercept	-5.471	-6.050	-3.665	-2.739	-2.534	-4.588
	(6.46)**	(6.27)**	(3.83)**	(2.56)*	(2.52)*	(5.41)**
Number of observations	2867	2867	2863	2421	2851	2862
Likelihood ratio test of independence between the selection and adequacy equations						
Chi2(1) test-statistic	1.48	2.31	1.95	1.42	0.70	1.97
p-value	0.22	0.13	0.16	0.23	0.40	0.16

Estimator is Heckman (selection) probit. Absolute value of t statistics in parentheses: * significant at 5%; ** significant at 1%
The corresponding selection regressions are shown in Table A2.

Table 6. Adequacy regressions with additional controls

	Subjective adequacy of:					
Isolation	**Food**	**Clothing**	**Housing**	**Schooling**	**Health care**	**Income**
Travel time to local markets (log)	-0.102	-0.073	-0.136	-0.231	-0.241	-0.039
	(2.43)*	(1.51)	(2.51)*	(4.22)**	(4.52)**	(0.77)
Urban population within 2 hrs travel time	1.056	1.042	1.833	1.253	1.652	1.053
	(3.88)**	(3.87)**	(5.31)**	(4.26)**	(4.89)**	(3.67)**
Population Density (per sqkm)	0.000	0.000	-0.001	0.000	0.000	0.000
	(1.70)	(1.03)	(2.18)*	(0.20)	(0.50)	(1.44)
Additional controls for isolation						
Travel time to nearest school (log)	-0.203	-0.006	0.093	-0.230	-0.026	-0.006
	(1.34)	(0.04)	(0.60)	(1.31)	(0.17)	(0.04)
Travel time to nearest health facility (log)	-0.142	-0.027	-0.051	-0.008	-0.027	-0.087
	(1.67)	(0.29)	(0.48)	(0.07)	(0.28)	(0.95)
Herfindhal index for food consumption	0.714	0.100	0.336	1.196	0.043	-0.027
	(1.03)	(0.15)	(0.47)	(1.30)	(0.06)	(0.05)
Herfindahl index for non-food consumption	2.211	0.785	2.885	1.050	2.795	2.236
	(2.47)*	(0.71)	(1.89)	(0.71)	(2.30)*	(2.04)*
Herfindahl index for durables consumption	-0.216	0.404	0.155	0.442	0.063	0.007
	(0.59)	(1.06)	(0.41)	(0.98)	(0.17)	(0.02)
Additional controls for subjective well-being						
Education of household head (log)	0.229	0.183	0.137	0.175	0.136	0.122
	(6.27)**	(4.29)**	(3.18)**	(4.21)**	(3.62)**	(3.14)**
One or more household members unemployed	-0.224	-0.329	-0.380	-0.240	-0.288	-0.433
	(3.03)**	(4.20)**	(4.75)**	(2.69)**	(3.77)**	(5.09)**
Days lost to illness over preceding year	-0.012	-0.006	-0.007	-0.008	-0.023	-0.015
	(2.71)**	(1.52)	(1.59)	(1.63)	(4.72)**	(2.96)**
Household hires permanent or casual labor	0.141	-0.003	0.210	-0.146	-0.122	-0.061
	(1.30)	(0.02)	(1.93)	(1.12)	(1.17)	(0.52)
Gini coef. of inequality in per capita consumption	0.330	0.596	0.190	1.179	0.112	-0.290
	(0.65)	(1.09)	(0.32)	(1.81)	(0.20)	(0.61)
Average rainfall in ward	0.000	0.000	0.000	0.000	0.000	0.000
	(0.86)	(1.11)	(1.95)	(2.93)**	(2.39)*	(0.21)
Standard deviation of rainfall in ward	0.001	0.000	0.001	0.000	0.000	0.000
	(2.15)*	(0.37)	(1.72)	(1.49)	(1.10)	(1.11)
In a district most affected by Maoist insurgency	0.490	0.223	0.193	0.293	0.355	-0.077
	(1.92)	(0.94)	(0.70)	(1.13)	(1.13)	(0.29)
In a district affected by Maoist insurgency	0.089	0.277	0.392	0.326	0.510	0.337
	(0.56)	(1.64)	(1.95)	(1.56)	(2.60)**	(1.93)
In a district little affected by Maoist insurgency	-0.041	-0.134	-0.239	-0.185	-0.131	-0.129
	(0.31)	(1.01)	(1.65)	(1.18)	(0.89)	(1.02)
Household consumption and controls			included but not shown			
Prices			included but not shown			
Regional and belt dummies			included but not shown			
Ethnicity dummies			included but not shown			
Intercept	-5.314	-6.784	-4.530	-2.947	-3.079	-4.355
	(5.50)**	(6.50)**	(4.29)**	(2.29)*	(2.83)**	(4.04)**
Number of observations	2867	2867	2863	2421	2851	2862
Likelihood ratio test of independence between the selection and satisfaction equations						
Chi2(1) test-statistic	12.88	0.01	0.70	0.00	0.47	0.12
p-value	0.00	0.92	0.40	1.00	0.49	0.73

Estimator is Heckman (selection) probit. Absolute value of t statistics in parentheses: * significant at 5%; ** significant at 1%
The corresponding selection regressions are not shown to save space.

Table 7. Adequacy regressions with average ward distance to nearest market

	Subjective adequacy of:					
Isolation	**Food**	**Clothing**	**Housing**	**Schooling**	**Health care**	**Income**
Average travel time to markets in ward	-0.160	-0.126	-0.136	-0.328	-0.333	-0.072
	(3.14)**	(2.16)*	(2.01)*	(5.24)**	(5.04)**	(1.17)
Other regressors	as in Table 6 but not shown here to save space					

Estimator is Heckman (selection) probit. Absolute value of t statistics in parentheses: * significant at 5%; ** significant at 1%
The corresponding selection regressions are not shown to save space.

Table 8. Instrumenting regression for consumption expenditures

Characteristics of household head	Coef.
Age of household head	0.033
	(7.20)**
Age of household head squared	0.000
	(4.66)**
Log(education of household head)	0.223
	(15.66)**
Instruments	
Education of father of head (log)	0.074
	(1.48)
Father employed in non-farm job	0.113
	(1.15)
Rainfall x father in non-farm job	0.000
	(2.17)*
St.dev. of rainfall x father in non-farm job	-0.001
	(2.67)**
Rainfall x father's education	0.000
	(0.64)
St.dev. of rainfall x father's education	0.000
	(0.50)
Ethnicity dummies	yes
Intercept	9.082
	(85.35)**
Number of observations	2862
R-squared	0.25
Joint F-test of the instruments	
Chi2 test-statistic	11.25
p-value	0.00

Estimator is OLS. Absolute value of t statistics in parentheses:
* significant at 5%; ** significant at 1%

Table 9. Adequacy regressions with instrumented consumption expenditures

	Subjective adequacy of:					
Isolation	**Food**	**Clothing**	**Housing**	**Schooling**	**Health care**	**Income**
Travel time to local markets (log)	-0.094	-0.070	-0.136	-0.221	-0.236	-0.038
	(2.32)*	(1.49)	(2.52)*	(4.09)**	(4.37)**	(0.75)
Urban population within 2 hrs travel time	1.010	1.012	1.842	1.200	1.625	1.041
	(3.87)**	(3.91)**	(5.33)**	(4.19)**	(4.86)**	(3.70)**
Population Density (per sqkm)	0.000	0.000	-0.001	0.000	0.000	0.000
	(1.79)	(1.10)	(2.14)*	(0.10)	(0.50)	(1.41)
Household consumption and controls						
Consumption expenditures (log)	1.710	1.789	1.298	1.769	0.777	1.399
	(4.94)**	(4.03)**	(2.91)**	(3.96)**	(1.44)	(3.26)**
Residuals from instrumenting regression	-1.272	-1.371	-1.150	-1.570	-0.664	-1.131
	(3.66)**	(3.08)**	(2.63)**	(3.54)**	(1.23)	(2.65)**
Other regressors	as in Table 6 but not shown to save space					
Likelihood ratio test of independence between the selection and satisfaction equations						
Chi2(1) test-statistic	23.62	2.66	0.22	1.83	1.27	1.23
p-value	0.00	0.10	0.64	0.18	0.26	0.27
Overidentification test for consumption expenditures						
Sargan statistic	3.65	1.79	2.92	3.18	2.32	1.98
p-value	0.60	0.88	0.71	0.67	0.80	0.85

Estimator is Heckman (selection) probit. Absolute value of t statistics in parentheses: * significant at 5%; ** significant at 1%
The corresponding selection regressions are not shown to save space. Overidentification tests are computed using OLS.

Table 10: Compensating variation of travel time to markets

	Moving from 75th to 25th distance percentile			At mean (1)	At 90% perc. (2)
	CV	(95% confidence interval)		CV	CV
Food consumption	13.0%	3.4%	22.8%	32.6%	3.9%
Clothing	10.3%	-0.1%	20.1%	26.2%	3.0%
Housing	22.5%	4.4%	40.7%	51.1%	6.9%
Children's schooling	22.0%	8.2%	35.8%	50.2%	6.8%
Health care	[20.2%-60.8%]	[13.4%-46.9%]	[27.0%-74.7%]	[46.9%-92.7%]	[6.2%-23.2%]
Total income	8.2%	-4.6%	21.1%	21.4%	2.4%
Weights equal to consumption share	[13.7%-15.7%]			[33.9%-38.1%]	[4.1%-4.7%]

All figures expressed in percentage of average consumption expenditures.

(1) Compensating variation generated by moving from the mean distance to the smallest recorded distance of 1 minute.

(2) Compensating variation generated by reducing distance by the same amount as in (1) but from the 90% percentile.

Table A1. House rental regression

House characteristics	*Coef.*	*t-stat.*	
Square footage of dwelling (log)	0.267	7.44	***
Square footage of plot (log)	-0.047	-0.29	
Number of rooms (log)	0.454	8.69	***
Share of kitchen in number of rooms	1.821	0.47	
Share of bathroom in number of rooms	2.111	0.55	
Share of bedrooms in number of rooms	2.122	0.55	
Share of living room in number of rooms	2.195	0.57	
Share of business room in number of rooms	2.713	0.71	
Share of mixed use room in number of rooms	1.585	0.41	
Share of other room in number of rooms	1.878	0.49	
Dummy if kitchen garden	0.053	1.27	
Construction material:			
Outside walls (omitted variable = cement or concrete)			
Mud cemented bricks or stones	-0.423	-6.06	***
Wood or branches	-0.368	-4.23	***
Other permanent material	-0.536	-5.73	***
Flooring (omitted variable = earth)			
Flooring material	0.323	4.94	***
Roof (omitted variable = thatch, wood or mud)			
Galvanized iron	0.467	7.49	***
Concrete or cement	0.498	5.48	***
Tiles or slates	0.418	7.45	***
Windows (omitted variable = no window coverings)			
Shutters	0.278	5.44	***
Screen or glass	0.549	5.97	***
Other	0.255	1.28	
Amenities:			
Dummy if electricity	0.297	5.66	***
Dummy if some sewerage system	0.230	3.09	***
Dummy if some garbage collection	0.408	4.97	***
Water (omitted variable = piped water)			
Covered well/hand pump	-0.065	-1.10	
Open well	0.015	0.15	
Other water source	-0.196	-2.53	**
Toilet (omitted variable = household flush)			
Household or communal latrine	-0.338	-4.78	***
No toilet facilities	-0.523	-7.47	***
District dummies	included but not shown		
Number of observations	5071		
R-squared	0.9607		

The dependent variable is the log of imputed rental value per month, in Nepalese rupees.

Table A2. Non-migration selection regressions

	Selection regression corresponding to subjective adequacy regression of:					
Characteristics of hhold head	**Food**	**Clothing**	**Housing**	**Schooling**	**Health care**	**Income**
Age of household head	-0.001	-0.002	0.000	0.012	-0.002	-0.001
	(0.08)	(0.18)	(0.01)	(1.03)	(0.15)	(0.07)
Age of household head squared	0.000	0.000	0.000	0.000	0.000	0.000
	(0.54)	(0.45)	(0.63)	(1.64)	(0.49)	(0.58)
Log(education of household head)	-0.110	-0.141	-0.133	-0.098	-0.129	-0.129
	(3.06)**	(4.38)**	(4.07)**	(2.68)**	(3.84)**	(4.04)**
Father employed in non-farm job	-0.055	-0.117	-0.091	-0.065	-0.086	-0.089
	(0.68)	(1.48)	(1.14)	(0.76)	(1.06)	(1.11)
Education of father of head (log)	-0.301	-0.307	-0.319	-0.329	-0.309	-0.316
	(6.61)**	(6.89)**	(7.61)**	(7.39)**	(7.22)**	(7.46)**
Inherited land in ha (log)	0.796	0.693	0.736	0.830	0.750	0.739
	(7.53)**	(5.49)**	(6.60)**	(6.95)**	(6.80)**	(6.71)**
Ethnicity dummies	yes	yes	yes	yes	yes	yes
Intercept	1.170	1.216	1.186	0.555	1.207	1.197
	(4.02)**	(4.30)**	(4.16)**	(1.75)	(4.23)**	(4.19)**
Number of observations	2867	2867	2863	2421	2851	2862

Estimator is Heckman (selection) probit. The second part of the regression is shown in Table 5.
The dependent variable is 1 if the head of household resides in birth ward, 0 otherwise.
Absolute value of t statistics in parentheses: * significant at 5%; ** significant at 1%