ISSN 1471-0498

DEPARTMENT OF ECONOMICS

DISCUSSION PAPER SERIES

ISOLATION AND SUBJECTIVE WELFARE

Marcel Fafchamps and Forhad Shilpi

Number 216

December 2004

Manor Road Building, Oxford OX1 3UQ

Isolation and Subjective Welfare*

Marcel Fafchamps
University of Oxford †

Forhad Shilpi
The World Bank‡

2004

Abstract

Using detailed geographical and household survey data from Nepal, this article investigates the relationship between isolation and subjective welfare. This is achieved by examining how distance to markets and proximity to large urban centers affect responses to questions about income and consumption adequacy. Controlling for migration, results show that isolation significantly reduce subjective assessments of income and consumption adequacy. Part of this effect can be attributed to lower access to public goods and to a reduction in the variety of consumption items. Compensating variation estimates suggest that the subjective cost of isolation is large. We also find strong evidence that Nepalese households cannot relocate costlessly out of their village of origin.

JEL codes: D60, I31, R20

Keywords: geographical isolation; consumption adequacy; distance to markets; subjective well-being

*We benefitted from comments from Mukesh Eswaran, 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.

†Department of Economics, University of Oxford, Manor Road, Oxford OX1 3UQ. Email: marcel.fafchamps@economics.ox.ac.uk. Fax: +44(0)1865-281447. Tel: +44(0)1865-281446.

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

1. Introduction

It is increasingly recognized that welfare varies dramatically between regions of the same country. This is particularly true since the methodological breakthrough of Elbers, Lanjouw & Lanjouw (2003) made it possible to compute spatial poverty maps for many developing countries (e.g. Ferreira, Lanjouw & Neri 2003, Alderman, Babita, Demombynes, Makhatha & Ozler 2002, Mistiaen, Ozler, Razafimanantena & Razafindravonona 2002). The existing evidence suggests that poverty concentrates in areas characterized by lack of infrastructure and poor access to markets, which are typical of geographically isolated communities (e.g. Jalan & Ravallion 2002, Ravallion & Jalan 1996, Ravallion & Jalan 1999, Ravallion & Wodon 1999). The effect of isolation on incomes works partly through its influence on the way people interact with markets and what they specialize in (e.g. Venables & Limao 2002, Fujita, Krugman & Venables 1999, Jacoby 2000, Fafchamps & Shilpi 2003, Fafchamps & Shilpi 2005). Isolation may also be correlated with certain forms of criminal activity (e.g. Fafchamps & Moser 2003, Fafchamps & Minten 2004) and, in the case of the country we study, with infiltration by guerilla groups.

While much has been written on the relationship between geographical location and objective measures of consumption and welfare, 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.[1] This paper revisits the question of the relationship between isolation and subjective welfare.

[1] Gugler (1991) provides some evidence that, in Nigeria, urban parents send their children to be raised in the surrounding countryside, suggesting that village is a better – or cheaper – place to raise their children.

At the heart of this question lies the issue of the spatial mobility of the labor force. If people are free to move and can do so costlessly, a simple spatial arbitrage argument predicts that utility should be equalized across locations. If this is the case, there is no point in studying the relationship between isolation and subjective well-being since, by assumption, there is none. This is, for instance, the approach taken by Rosen (1979) and Roback (1982) who proposed a hedonic pricing approach to calculate the implicit quality of life of various US locations. Using the assumption that utility is equalized across locations, they were able to use price and income differences across locations to infer the welfare cost or benefit of location-specific attributes.

In contrast, if people cannot move or can only do so at a cost, utility need not be equalized across space and a relationship between isolation and utility can arise and be studied. Isolation may raise subjective satisfaction directly if people value social cohesion, closeness to nature, and lack of congestion. It may also reduce it because of poor access to public services such as schools, hospitals and road infrastructure and by a reduced choice of consumer goods and services.

The purpose of this paper is to take a fresh look at this issue by estimating 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 making any assumption about spatial mobility.

The starting point of our empirical specification is a standard utility maximization model in which isolation affects 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.[2]

The consumption and income adequacy questions are intended to measure subjective satisfaction with the standards of living of the household. They are simple to understand, even to illiterate households, and have been used in many surveys. The reader may nevertheless wonder whether responses to these questions – and more generally to subjective well-being assessment questions – can be taken as measures of utility. While there is no denying that answers to such questions contain a lot of 'noise', there is increasing evidence that they are valid measures of subjective welfare. Frey & Stutzer (2002) and Diener, Suh, Lucas & Smith (1999), for instance, document various psychology experiments demonstrating that they are correlated with cat scans, clinical depression, and suicide rates. A strong empirical association has been found between income and answers to the subjective well-being question (e.g. Easterlin 1974, Easterlin 1995, Easterlin 2001, Blanchflower & Oswald 2004, Di Tella, MacCullock & Oswald 2001, Frey & Stutzer 2002, Diener et al. 1999).[3] Recent evidence of this relationship is, for instance, provided by Kingdon & Knight (2003) and Lokshin, Umapathi &

[2] 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.

[3] The relationship between well-being and income appears to break down over time. For instance, over the last 50 years the US has experienced no change in subjective well-being despite a 6-fold increase in income per head (e.g. Easterlin 1974, Frey & Stutzer 2002). Since we only have cross-section data, this does not affect our analysis. Inter-country comparisons have further shown that diffentials in subjective well-being across countries are not commensurate to differentials in GDP per head, especially at higher levels of income (Inglehart & Klingemann 2000). The breakdown of the relationship between well-being and income over time is usually blamed on two interrelated phenomena: habituation – what Brickman & Campbell (1971) call the 'hedonic treadmill'; and rivalry – what many authors label the 'relative income' effect. The applicability of these ideas to Nepal is examined by Fafchamps & Shilpi (2004).

Paternostro (2003) for China and Madagascar, respectively. Using data on Nepal and Jamaica, Pradhan & Ravallion (2000) demonstrate that responses to the questions on subjective consumption adequacy are strongly correlated with more objectively measured poverty indicators.

Answers to subjective welfare assessment questions are correlated with other factors as well. Drawing from experimental and survey results, responses to subjective well-being and consumption adequacy questions have been shown to be also affected by psychological factors such as mood and temperament (e.g. Diener et al. 1999, Kahneman, Diener & Schwarz 1999). They are also influenced by unemployment, inflation, gender, race, ethnicity and a number of institutional factors such as democracy and governance (e.g. Easterlin 1974, Easterlin 1995, Easterlin 2001, Di Tella, MacCullock & Oswald 2001, Ravallion & Lokshin 2001, Frey & Stutzer 2000, Diener & Oishi 2000, Diener, Diener & Diener 1995, Inglehart 1990, Inglehart & Klingemann 2000, Kingdon & Knight 2004). In the empirical analysis, we control for as many of these factors as possible to ensure that the relationship we uncover is not due to omitted variable bias.

Our econometric investigation leads to a robust finding: isolation has a negative welfare effect. Average village isolation (measured by the average village level travel time to the nearest market) has a significant and large negative effect on subjective well-being while proximity to urban centers increases it significantly. Household isolation within the village (measured by household-specific travel time to nearest market) also has a significant negative effect on certain dimensions of consumption. These results obtain after we control for consumption expenditures and prices, suggesting that isolation has a welfare cost over and above its effect on monetary consumption. Controlling for household mobility and adding many controls leaves results unchanged.

We quantify the subjective welfare loss associated with isolation. We find that the compensating variation of isolation is very large, particularly with respect to housing, schooling and

health care. This suggests that isolation has a direct subjective welfare cost that can potentially be remedied by improved road infrastructure and the provision of public services.

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 and subjective cost of isolation is estimated in Section 5. Section 6 concludes the paper.

2. Conceptual framework

We are interested in finding out how isolation affects household welfare. For the purpose of this paper, we define isolation as distance from urban centers: households located at a large distance d from the nearest urban center are deemed to be more isolated than households located closer. Since the purpose of this section is to motivate the econometric analysis, we keep the model as simple and general as possible.

2.1. Utility and isolation

The starting point for our analysis is standard utility theory (e.g. Ravallion & Lokshin 2001, Lokshin, Umapathi & Paternostro 2003). Households are assumed to derive satisfaction or utility from consumption c. Isolation d may affect utility through its effect on total consumption expenditures $y(d)$ and prices $p(d)$, on the range $N(d)$ of available goods and services, and on public goods and externalities $G(d)$.[4] We assume that public goods and externalities are taken

[4] Without loss of generality, the utility households may derive directly from living away from others can be regarded as part of $G(d)$: absence of congestion and closeness to nature are treated as a public good consumed by the household.

as given by the household. The household's problem can thus be written:

$$\max_{c} U(c, G(d)) \text{ subject to } y(d) = \sum_{i=1}^{N(d)} p_i(d) c_i \tag{2.1}$$

The solution to the above problem is an indirect utility function of the form $V(y(d), p(d); N(d), G(d))$. As usual, we have with $\partial V/\partial y > 0$ (non-satiation) and $\partial V/\partial p_i \leq 0$ for all $i \in N(d)$. In addition, we also have $\partial V/\partial N \geq 0$ from Le Chatelier principle: optimization problems in which the number of available goods is small can be seen as constrained versions of optimization problems with more goods to choose from. This condition is satisfied, for instance, when utility is of the Cobb-Douglas or CES form. We also normalize $G(d)$ so that $\partial V/\partial G > 0$.

We have little a priori expectations regarding the relationship between prices, nominal income, and distance. This is because the effect of distance on prices depends on whether goods are exported or imported: if a good is exported from location k to location m, arbitrage dictates that the price in k should be below that in m to compensate for transport costs, and vice versa if the good is imported. We nevertheless expect real consumption in an isolated location k to be decreasing in distance because gains from trade and specialization are not achieved. Fafchamps & Shilpi (2005) provide indirect evidence that the range of goods and services produced in a location increase with its proximity to urban centers.

Regarding the availability and variety of goods, we expect N to be decreasing in distance: because gains from specialization increase with the

extent of the market, cities produce more variety of goods and services. Isolated areas must import this variety from urban centers. It is not profitable to transport the full range of goods to isolated areas whenever there are fixed transport and transactions costs per variety, such as ordering or menu costs. Reduction in variety is expected to be particularly strong for services because they typically have to be provided locally.

We have no a priori regarding the possible relationship between G and distance. Congestion concerns militate for a positive relationship, with a better quality of life in a less urban setting. Access to public services such as schools and hospitals, however, militate for a negative relationship. Fear of crime, wildlife, and natural elements may also fuel a negative relationship (Fafchamps & Moser 2003). Other factors may operate in the opposite direction. It has been argued, for instance, that people assess their consumption needs by comparing themselves to others (e.g. Duesenberry 1949, Frey & Stutzer 2000, Layard 2002, Curtis & Eswaran 2003, Kingdon & Knight 2004). By exposing households to outside influences, proximity to urban areas may reduce their satisfaction level by raising expectations and making people feel that their standard of living is inadequate (Fafchamps & Shilpi 2004).

If the household produces part of what it consumes, the optimization problem takes the more general form:

$$\max_{c,q} U(c, G(d)) \text{ subject to } \pi(p; d, A(d)) + y(d) = \sum_{i=1}^{N(d)} p_i(d)c_i \equiv E(d) \tag{2.2}$$

where $\pi(p; d)$ denotes the household's profit function and $A(d)$ stands for the assets available to the household (e.g., land). Because land located far from cities has a lower value (Jacoby 2000), it is possible – though by no mean necessary – that land endowments increase with distance. In contrast, we expect human capital to be higher near urban centers because schooling is more readily available and there are more jobs for educated workers. The solution to (2.2) is of the form:

$$V = V(y(d) + \pi(p; d, A(d)), p(d); N(d), G(d)) \tag{2.3}$$

Because prices enter the profit function, it follows that the sign of dV/dp_i is indeterminate whenever price p_i enters the profit function, i.e., whenever good i is produced or used as production

input by the household.

The purpose of this paper is to estimate an empirically implementable version of equation (2.3) directly, controlling for total consumption expenditures X by household j, prices p_j in the location where j resides, distance d_j between j and the nearest market, and household taste shifters Z_j:

$$V_j = V(X_j, p_j, d_j, Z_j) \tag{2.4}$$

Measures of V_j are obtained from subjective rankings provided by respondent households.

If we do not control for N or G directly, the coefficient of distance captures the effect of both. Differences in the effect of distance on our six measures of subjective satisfaction provides useful albeit impressionistic information on the relative importance of N and G in the relationship between distance and utility. We expect satisfaction with access to health care and schooling to depend on public provision of such services locally. In contrast, satisfaction with food and clothing is expected to vary with the variety of goods for sale on local markets. Satisfaction with housing quality may also depend on the local availability of construction materials (e.g., iron roofing, doors, sanitary and electrical equipment). Because job specialization and economic diversification are higher in and around urban centers (Fafchamps & Shilpi 2005), we also expect consumption variety to increase with the size of nearby urban population.

To summarize, we expect satisfaction with health care and schooling to depend on G but not on N while the opposite is true for satisfaction with food, clothing and housing. At the limit, if all the welfare cost of isolation is due to N, we expect isolation to affect responses to the food, clothing, and housing but not the schooling and health care satisfaction questions. A contrario, if all is due to G, distance should affect only responses to the schooling and health care questions.

Satisfaction with total income should, in principle, combine all the effects of distance to-

gether. It is, however, conceivable that respondents regard monetary issues as separate from problems of access to variety (N) or to public service (G). Someone may, for instance, answer that his income is adequate but complain that he cannot buy the clothing he desires because it is not available locally, or that his children cannot attend school because it is too far.[5] If this is the case, satisfaction with monetary income would vary less with distance since it would not include the effects of N and G.

Equation (2.4) can be used to compute the welfare cost of isolation C_j by applying the standard definition for compensating (or equivalent) variation:

$$V(X_j, p_j, 0, Z_j) = V(X_j - C_j, p_j, d_j, Z_j) \tag{2.5}$$

where C_j represents the monetary transfer household j would be willing to pay to eliminate isolation completely. The more isolated household j is, the more we expect it to be willing to pay to set $d_j = 0$. How equation (2.5) can be used in practice to estimate C_j is discussed in Section 5.

2.2. The choice of location

So far we have reasoned as if the location in which people is exogenously determine. We now endogenize the choice of location.[6] Let function $X(d_j, Z_j)$ capture the dependence of income and hence expenditures on distance and household attributes. Similarly, let $p(d_j)$ denote the dependence of prices on distance. We assume that heads of household choose where to locate,

[5] To an economist it would seem that a sufficiently high income would enable someone to overcome insufficient access (e.g., by paying a private tutor or paying someone to buy the clothes in town). According to this reasoning, insufficient access is ultimately an income problem. This is probably not how most respondents see it. Having a private doctor or tutor, for instance, is not within their frame of reference. It is therefore likely that, for most respondents, issues of access are distinct from issues of monetary income.

[6] To keep things simple, in the presentation here we proceed as if distance d_j is the only characteristic across which locations differ. In the empirical analysis, we generalize this by allowing various regional and climatic effects to matter as well.

that is, they choose distance d_j by solving an optimization problem of the form:

$$\max_{d_j} V_j = V(X(d_j, Z_j), p(d_j), d_j, Z_j) \tag{2.6}$$

Suppose for a moment that Z_j is the same for all households. If they locate freely and costlessly, all choose the same optimum distance d_j. For individuals to accept to live at different distances, incomes and prices must adjust in equilibrium to make household indifferent between these distances. Following Rosen (1979) and Roback (1982), a pure arbitrage argument therefore predicts that utility is equalized across locations although income and prices need not. In this context, incomes differ across location but not welfare because differences in income are purely compensated by differences in prices and location characteristics. Observing that people living in isolated areas are poorer does not imply the need for policy intervention.

The above reasoning generates a simple but powerful test. Suppose we observe a measure of V_j. If people can move freely, then there should be no statistical relationship between this measure and d_j. It follows that if V_j varies systematically with d_j, this means that people cannot move freely and costlessly – and thus that there might be a rationale for policy intervention to redress welfare differences between locations.

This conclusion rests crucially on the assumption that Z_j does not vary across individuals. Let us now assume that this is not the case. In particular, assume that certain people are anti-social and prefer isolation while others do not, i.e., that the cross-partial derivative $V_{dZ} \neq 0$. Without going into details we expect individuals with different preferences to sort themselves by distance: in equilibrium, antisocial people should take isolated locations while social people select locations that favor human interaction. In such a world, people living in isolated villages would simply be there because they love isolation.

If we can control for Z_j, individual heterogeneity does not affect the validity of the test.

People with the same Z_j should be indifferent between locations and thus, conditional on Z_j, we should not observe any systematic relationship between V_j and d_j if people can move freely and costlessly. Things are different if (part of) the vector Z_j is unobserved. To illustrate this point, suppose that some people are grumpy, i.e., that they report a lower level of V_j or, in our case, of consumption adequacy. Further suppose that grumpy people are also antisocial, i.e., that they prefer isolation. As grumpy people choose to locate in isolated areas, this generates a negative correlation between distance and reported consumption adequacy. Since we do not observe who is grumpy, when we regress our measure of V_j on d_j and the controls we obtain a spurious correlation between d_j and the error term.

Things are different if people can only move at a cost. If the cost is so high that nobody moves, we are back to the exogenous case: where people live is fully determined by where they happen to be born. In this case, regressing V_j on d_j captures the welfare cost of isolation. At intermediate values of the relocation cost, only a certain proportion of people move. For those who do not, location remains exogenously determined. Only those households who move choose the place where they live. In contrast, non moving households do not choose their d_j – except to the extent that they choose not to move.

Focusing on non-movers therefore eliminates the endogeneity of d_j. The possibility of a selection bias remains, however. This is because, for a given cost of moving, households are more likely to move if the gain from moving is large. To illustrate how this would affect inference, continue to assume that individuals who like isolation are grumpy. People who dislike isolation benefit the most from moving closer to the market or the city, and this benefit grows the further away from the city they are born. This means that grumpy people are over-represented in the population of non-moving individuals, and that this over-representation increases with distance. In short, if we regress V_j on d_j for non-movers only, we have to correct for the possibility of

selection bias.

Given the circumstances prevailing in the studied country, it is reasonable to assume that heads of households make the decision of where to locate at the beginning of their productive life. We observe a representative sample of the population. Some heads of household have moved and are surveyed in a village or town other than their birthplace; others have not. Let $M_j = 0$ is the head of household lives in his birth village, and 1 otherwise. This variable can be used to estimate a selection equation. We just have to be careful only to use regressors that were determined before the decision to relocate was taken. Let H_j be such a vector of regressors thought to affect the decision to migrate out of one's birth place. Examples of variables that can be included in H_j are education and parental background.[7] Date of birth can also be included to reflect changes in economic opportunities over time.

To summarize, our estimated model has the form:

$$\begin{aligned} V_j &= \alpha + \theta \log d_j + \beta \log X_j + \gamma \log p_j + \lambda Z_j + u_j \text{ if } M_j = 0 \\ M_j &= 1 \text{ if } M_j^* = \rho H_j + v_j \geq 0 \\ &= 0 \text{ if } M_j^* = \rho H_j + v_j < 0 \end{aligned} \tag{2.7}$$

where M_j^* is denotes the usual latent variable associated with the relocation decision. The above model is estimated using standard selection correction techniques. For instance, if V_j only takes two values, model (2.7) can be estimated using Heckman's maximum likelihood probit estimator with selection correction.

[7] Here we implicitly assume that parents choose the level of education of their children. Given the very low levels of schooling achieved by our sample population, this is a reasonable assumption.

3. The data

The data we use come from the 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 where 'adequacy' was defined as the level of consumption which is no more or no less than what the respondent (household head) considered to be the minimum needs of the family. The household head was asked to indicate on a three point scale whether household's consumption of food, clothing and housing, and household's access to children's schooling and health care, and total income earned during the last month were adequate or not.[8]

Responses to these questions are summarized in Table 1. The overall dissatisfaction of household heads about their life is quite striking. About 69 percent of the households have less than adequate income according to the self-assessment of the household head. 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.[9] For the other five questions, more than half of the household heads feel that their consumption and/or access levels are less than what they consider to be the minimum needs of the household. Only a small proportion of households report their income or consumption to be more than adequate. The responses to these qualitative questions constitute the dependent variables used in our analysis.[10]

[8] To be more precise, in the survey, the household head was asked the following question: Concerning your family's food consumption over the last 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].

[9] The incidence of poverty measured by the head count ratio is estimated to be around 42 percent (World Bank, 1999).

[10] In econometric analysis, responses are transformed into a numerical scale – 1 for not adequate, 2 for adequate and 3 for more than adequate.

The survey collected information on a whole range of household level variables including size and composition, and asset ownership. Household characteristics and ward-level variables 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. 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. We see that, on average, surveyed households live more than two hours of travel time from a market, the maximum value being 40 hours. The median is around 1 hour.

The average total annual consumption (non-durables and durables) of each household is reported in US$ to facilitate interpretation.[11] The total value of assets is reported as well. This includes land, livestock, agricultural equipment, and financial assets. As is customary with wealth data, we note that the wealth distribution is quite unequal (high standard deviation) and highly skewed, with the median representing around one fourth of the mean. The average age of the household head is 45 years and the average household size is 5.6 individuals.

Table 2 reports the distances between the household and the nearest school and health facility. These variables are used to proxy for the provision of public services. As we can see, the average household is located around 20 minutes from the nearest school and one hour from the nearest health facility. There is, however, a lot of variation across households. Over the year preceding the survey, households lost on average 2.3 days of usual activity due to illness.

On average, adult females represent a quarter of household members. Some 13.5% of sampled

[11] 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.

households are headed by women. 18% of surveyed households have at least one unemployed member. Only 20% of surveyed households live outside the ward in which the head of household was born. Put differently, 80% of households live in the ward where the head of household was born. This suggests that the Nepalese population is much less mobile than, say, the US population. With so little movement of population, we should not be overly surprised if the assumption that people can move freely and costlessly is rejected for Nepal.

Parental 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. These variables are used to predict migration out of one's birth ward. We expect individuals raised by a better educated father working in a non-farm job to be more likely to migrate to town. In contrast, we expect household heads who inherit a lot of land from their father to be less likely to leave the countryside.

Ethnicity dummies are reported as well. Ethnicity may affect well-being by its effect on social status. It may also affect people's income opportunities and hence spatial mobility. As is immediately apparent from the list, ethnicity in Nepal is defined partly by caste and partly by religion. Brahmins and Chhetry are Hindu upper castes categories. The Newar are a tribal group with its own caste system. Together, these three groups account for 45% of the surveyed households. There are a number of other tribal groups as well, and various Hindu lower caste groups. Some ethnic groups are defined on the basis of religion instead. Other groups account for another 17% of the sample, suggesting the presence of a lot of ethnic diversity in the country. This is hardly surprising given that the mountainous nature of the terrain does not favor human interaction and has isolated people for much of the country's history.[12]

Next we list the ward-level variables used in this analysis. Using detailed information on

[12] The first road to Katmandu was built in 1929. Many districts of the country still have no road.

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 living within 2 hours of travel distance from the ward. Population figures come from the 1991 census. This measure, which we call 'proximity to urban population', is our second isolation measure.

The survey collected information on rice prices at the household level. 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 village or ward face the same prices. Consequently, in our analysis we use instead the median rice price in each ward. The median wage rate in the ward is similarly computed from responses of individual household members about wage rates from wage employment in agriculture and non-agriculture. We also report Gini coefficients for consumption per capita computed for each ward. It is indeed widely thought that inequality affects subjective well-being negatively. Consequently, wards that are more unequal may be less satisfied. To this extent that inequality is stronger in urban areas, this could generate an omitted variable bias.

As a way of capturing the impact of variety N, we construct for each ward three indices of product 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:[13]

[13] 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.

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

This index gives a rough idea of what is available for sale in the ward or nearby: the higher its value, the more concentrated spending is on a small number of categories, and the less diversified ward consumption is.[14] As expected, index values reported in Table 2 show more concentration in durable expenditures and less in foodstuffs.

Two types of location variables are used in the econometric analysis to control for various climatic and economic factors: ecological belt dummies and regional dummies. Ecological belt dummies basically 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 the majority 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.

The Table also reports areas most affected by the Maoist insurgency as of June 2000. At the time the NLSS was conducted, there had been a lull in the Maoist uprising in the early 1990's. Guerilla activity was resumed in early 1996, as the survey was being completed in the field. In the 1990's police stations were the main target of the insurgency. We expect the insurgency to have had minimal impact on responses to survey questions. But in case they did, we report in Table 2 insurgency incidence figures based on a classification of the Nepalese police.[15] Some 12.5% of the surveyed households resided in areas that were affected or most affected by the

[14] $N^J = 1$ 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.

[15] The Kathmandu Post of June 15, 2000 reports that the Nepalese police have classified districts in terms of severity of Maoist insurgency as follows: 1. most affected: Jajarkot, Rolpa, Rukum and Salyan; 2. affected: Dolkha, Gorkha, Jumla, Kavre, Khotang, Pyuthan, Ramechhap, Sindhuli, Sindhupalchowk; 3. little affected: Arghakhanchi, Bardiya, Bhojpur, Dailekh, Dang, Dhading, Dolpa, Gulmi, Khotang, Lalitpur, Lamjung, Makwanpur, Nuwakot, Okhaldhunga, Surkhet, Syangja, Tanahu, and Udayapur. The rest are not affected.

insurgency in the late 1990's. These districts tend to be far from urban centers. At the bottom of the Table we report rainfall information. It is used as instrument for consumption in a way that is discussed more in detail in the empirical section.

4. Econometric results

4.1. Non-parametric analysis

We now turn to the econometric analysis. We begin with simple non-parametric regressions of answers to income and consumption adequacy questions on (log) distance to the market. 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.

It is immediately apparent that satisfaction with income and consumption falls dramatically and significantly with distance from markets. Virtually identical figures obtain if we only use non-migrant households. The relationship between distance and satisfaction is monotonic and basically linear, except at high market distances for which the small number of observations does not allow precise estimation. This constitutes preliminary evidence that the assumption that households can relocate freely and costlessly should be rejected.

Evidence from univariate regressions can be misleading, however. We are particularly concerned about a possible relationship between expenditures and distance to market: if consumption is lower in isolated households, this could explain their lower reported satisfaction level. Indeed if we regress consumption expenditures on distance, we find a strong negative relationship. This in turn could be misleading as well if isolated households are smaller or face lower prices. To investigate the relationship between distance to market and consumption expendi-

tures, we therefore estimate a partial non-parametric regression of the form:

$$\log X_j = \beta B_j + \varphi(d_j) + \varepsilon_j$$

where B_j is a vector of controls for household j and $\varphi(.)$ is an arbitrary smooth function. The purpose of the controls B_j is to correct for price differentials and for household composition, so as to de facto compare consumption per adult equivalent in real terms.[16] We also include age as it may affect the consumption needs of the household. Vector B_j includes ward median wage, ward median rice price, household size and composition, age and age squared of the head, regional dummies, and a female head dummy. The composition of the household is captured by the share of women, young children, youth, and elderly members in the household.[17] Estimated function $\varphi(d_j)$ is depicted in Figure 2. Results confirm the existence of a strong negative relationship between distance and consumption expenditures, even after we control for prices and household composition.

4.2. Multivariate analysis

To investigate the separate welfare cost of isolation, we now turn to a multivariate approach in which we control for consumption expenditures. We estimate an equation (2.4) of the form:

$$V_j^h = \alpha + \theta^h \log d_j + \beta^h \log X_j + \gamma^h \log p_j + \lambda^h Z_j + u_j^h \tag{4.1}$$

where V_j^h denotes the satisfaction ranking reported by the head of household for one of the income and the five consumption categories h discussed earlier. Consumption expenditures are

[16] We could have simply divided total expenditures by the number of household members, yielding consumption per head. But doing so may bias results due to economies of size in household production (e.g. Deaton & Paxson 1998, Fafchamps & Quisumbing 2003). Our approach is more general.

[17] Adult males are the omitted category.

included as well. Controls B_j are included in Z_j to correct for differences in prices and 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 it can derive from a given level of consumption expenditures. 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. To account for this possibility, we include the share of various age/sex groups in the regression.

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, Fafchamps 2003).

Multivariate regression results are presented in Table 3. Since dependent variables can take three ranked values, ordered probit is used as estimator. In this first set of results, we do not yet control for self-selection; this is done in the next section. All regressions show a negative effect of distance to the nearest market d_j on subjective well-being. The effect is strong and significant in

all regressions. Our second measure of isolation, urban population within two-hour travel time from the survey ward, is positive in all six regressions – and significant in five. This means that, after controlling for income and distance, subjective well-being is higher in households located in or nearby large urban centers. It is not just distance to the local market that matters, but also the size of the urban population in that market and the towns located nearby. Taken together, these results are a strong indication that subjective well-being is negatively affected by 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 effects. The magnitude of the distance coefficient is largest for health care and schooling, suggesting that G is 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 is consistent with the idea that answers to the income satisfaction question do not include respondents' assessment of N and G. We also find strong urban size effects for food and clothing consumption, consistent with the increase of variety in and around large urban centers. Taken together, the results suggest that both N and G are large components of the welfare cost of isolation. We revisit this issue in the next section.

Turning to other regressors, we find a positive and almost always significant effect of consumption expenditures on answers to income and consumption adequacy questions. Except for health care, the effect is highly significant. Household assets have a significant positive effect in all six regressions while household size has a significant negative effect. These results are consistent with the utility model. Price effects are in general ambiguous. Our two measures of village-level prices have a negative and significant effect on food consumption, but the reverse is true for health. In other regressions, price effects are mostly non-significant. As mentioned in Section 2, this could be because prices affect households as producers as well as consumers.

Several of the controls are significant with the anticipated sign. In most cases, the magnitude of the household size coefficient is commensurate with the consumption coefficient, suggesting that results would not change much if we simply divided consumption by the number of household members instead of entering both regressors independently. Household composition regressors are occasionally significant. Households with a larger proportion of children and youths tend to report a lower satisfaction level, especially in housing, clothing, and total income. The age of the head is only significant in the schooling case, probably because the age of the household heads determines whether he or she has – or is considering having – school age children. The female head dummy is negative in all regressions but never significant.

We also find strong regional effects. 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 the isolation idea: the steeper the terrain, the more arduous travelling to the market becomes and the less likely travel is to take place on motorized vehicles. The exception for health care may capture a deliberate government effort to provide better health care coverage in the Hills and Mountain zones. We also find strong differences along the East-West gradient, with Eastern and Western provinces more satisfied on average than those residing in the Central zone. This may capture proximity to India.[18]

4.3. Correcting for self-selection

In Section 2 we discussed the idea that, if households freely and costlessly choose where to live, any observed relationship between isolation and expressed satisfaction with consumption would reflect self-selection of 'grumpy' individuals in isolated areas. Given that 80% of the surveyed heads of household reside in their birth ward, it is highly likely that financial and psychological

[18] Nepal neighbors India on its Eastern and Western borders. Better road infrastructure in India may enable products to penetrate more easily in the Nepalese Hills and Mountain regions.

costs restrict population movements. It nevertheless remains conceivable (although improbable) that the results presented in Table 3 are due to a strong self-selection bias.

To investigate this possibility, we reestimate equation (4.1) using only non-migrant households. In addition, we correct for self-selection into migrant status as described in model (2.7). Migrant households are those who were surveyed in a ward other than the birth ward of the head. The regressors used in the selection equation are variables thought to affect the decision to migrate out of one's birth place, such as education and parental background. 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 as well if, as is likely, certain groups have better networks with migrant populations elsewhere (e.g. Seddon, Adhikari & Gurung 1999, Munshi 2003).

In order to estimate model (2.7) using the bivariate probit selection estimator of Heckman, 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 not too large.

Regression results are presented in Table 4 for V_j^h and in Table 5 for the selection regressions. Using a likelihood test, the absence of correlation between the errors in the selection and satisfaction regressions is rejected for the food and clothing regressions – and nearly rejected for other regressions except health care. A selection correction is thus appropriate. As is clear from Table 5, our results are basically unchanged: distance is significantly negative in all regressions, and larger in magnitude for schooling and health care. Consumption similarly remains positive except for health care, where it is still non-significant. Other qualitative results remain unchanged. Thus, although a selection correction is shown to be appropriate in this case,

self-selection does not appear to be responsible for our results.

Turning to Table 5, we find 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 1999). We also find that migrating out is more likely if the head's father was 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. Ethnic dummies are also significant, with those belonging to the Brahmin caste and to the Magar and Tharu tribes more likely to migrate.

4.4. Adding more controls

Next, we include in the regressions variables that proxy for N and G. As proxy for G, we include the (log of the) distances from the household to the nearest school and health facility. These measures do not account for likely differences in quality and breadth of service between different locations, on which we have no information. 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.

We also add a number of regressors to minimize the risk of omitted variable bias. In particular, we add the education of the household head. Education has indeed been shown to influence responses to subjective well-being questions Diener et al. (1999). Unemployment and illness are included into the regression 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 is included as well, together with the ward-specific standard deviation of rainfall in the year of the survey. These variables are included to capture possible effects of climate

on residents' mood. The Maoist insurrection may also affect people's mood, which is why we include our insurgency dummies.[19] 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.[20] For the same reason, we also include all ethnic dummies.

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.[21] Except for the income adequacy regression, our main results are basically unchanged: distance remains negative and significant while the urban population variable remains positive and significant. This demonstrates that the relationship between isolation and subjective adequacy assessments survives the elimination of many potential sources of omitted variable bias. In the income adequacy regression, while urban proximity remains significant, distance to market falls beneath standard levels of significance. This is consistent with variety N and access to public services G looming large in the welfare cost of isolation.

In each of the regressions, at least one of the variety indices N^J is significant with the anticipate positive sign. We take this as evidence that variety is valued by Nepalese households. The effect of the inclusion of these variables on the distance coefficients is minimal, however, suggesting that the N^J indices alone does not exhaust the effect of isolation. Distance to the nearest school and health facility have the anticipated negative sign in most regression but they

[19] 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.

[20] 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.

[21] 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. It therefore appears that the inclusion of additional regressors eliminates the self-selection bias, except in one regression.

are both significant only in the food regression. Distance to school is significant in the schooling regression. This suggests that other dimensions of local public service provision matter as well, such as the quality of the school or health facility and the availability of drugs and teachers, for which we do not have data.

Regarding other controls, since their inclusion is purely to eliminate possible sources of omitted variable bias, they need not be discussed in details. 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, 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 only one regression. 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.

Tables 4 and 6 control for self-selection across wards. The reader may nevertheless worry that placement within the ward remains endogenous and thus may bias results. To investigate this possibility, we reestimate Table 6 replacing individual distance by the ward average.[22] To save space, we show in Table 7 the regression results for the distance coefficients only. We see that

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

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, however, gets in the way of precise estimation. The results, not shown 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. In housing, both ward median and individual distance are negative but only individual distance is significant. In the total income regression, both variables are individually non-significant although they remain jointly significant ($\chi^2 = 5.10$, p-value$= 0.08$). Taken together, this evidence suggests that what matters most is isolation of the ward itself, not relative isolation of the individual within the ward. We take this to imply that endogenous placement within the ward is highly unlikely to be responsible for our results.

To close this section, we briefly discuss the issue of response bias. Social scientists indeed often voice their concern about the validity of statistical analysis in terms of response bias. Although all survey data are susceptible to response bias, the subjective nature of the consumption adequacy questions raises additional concerns. In econometric terms, response bias can lead to inconsistent estimates if it is correlated with the regressor of interest. In our case, for response bias to cause the observed relationship between isolation and consumption adequacy , it would have to be true that urban dwellers respond on average more positively to the consumption adequacy question than more isolated rural dwellers even though both are, in fact, equally satisfied.

There are several reasons why such an occurrence is highly unlikely. First, research indicates that people answer subjective well-being questions by comparing themselves to a reference point (e.g. Diener, Diener & Diener 1995, Diener & Oishi 2000). This reference point is partly affected by what other similar people consume (e.g. Frey & Stutzer 2002, Kahneman, Diener &

Schwarz 1999, Layard 2002). Isolated residents of rural Nepal are much less exposed to imported standards of living than their urban counterparts – and for this reason they are *more* likely to be satisfied than urban dwellers. Second, it has been shown that mood affects responses to subjective well-being questions (Diener & Biswas-Diener 2000). Having to respond to survey questions may affect people's mood negatively. If rural respondents are more affected by the presence of enumerators than urban dwellers, this could generate the observed negative relationship between distance and subjective adequacy assessments. In our experience, however, if anything it is the urban households who resent surveys more because the opportunity cost of time is higher in urban areas. Furthermore, for many rural inhabitants of poor countries answering survey questions is a totally unwelcome distraction. For all these reasons, we find it extremely unlikely that our results be an artifact of response bias.

5. Estimating the welfare cost of isolation

In this last Section, we use equation (4.1) to obtain a money-metric estimate of the welfare cost of isolation. Such a measure is useful because it gives an order of magnitude to the subjective cost associated with isolation.[23] Estimation is complicated by the fact that we do not have a single estimate of V_j but several partial rankings V_j^h. In order to make things manageable, let us postulate that:

$$V_j = \sum_{h=1}^{H} \gamma^h V_j^h + \varepsilon_j \tag{5.1}$$

Since utility is define only up to a monotonic transformation, the weights γ^h of the different rankings in total utility must be normalized in some appropriate way. Here we require that

[23] Although it may be tempting to use the figures reported here as basis for policy – e.g., to estimate the welfare gain from building a road – it would be hazardous to do so. This is because our analysis does not tell us why isolation decreaes welfare, i.e., whether it is because of isolation by itself, access to health and sanitation, access to markets, or other public services. These issues deserve more research. Also our analysis does not compute the income gain one may expect from building a road.

$\sum_{h=1}^{H} \gamma^h = 1$. From the previous section, we have estimates of the relationship between partial rankings V_j^h, distance d_j, and total expenditures X_j. Supposing we knew the γ^h's, we could use equation (4.1) in combination with (5.1) and (2.5) to compute C_j as the solution to:

$$\sum_{h=1}^{H} \gamma^h \left[\alpha^h + \beta^h \log X_j + \theta^h \log d^{\min} + u_j^h\right] + \varepsilon_j = \sum_{h=1}^{H} \gamma^h \left[\alpha^h + \beta^h \log(X_j - C_j) + \theta^h \log d_j + u_j^h\right] + \varepsilon_j \tag{5.2}$$

from which we obtain:

$$\log(1 - c_j) = \log \frac{X_j - C_j}{X_j} = \frac{\sum_{h=1}^{H} \gamma^h \theta^h (\log d^{\min} - \log d_j)}{\sum_{h=1}^{H} \gamma^h \beta^h} \tag{5.3}$$

where $c_j \equiv C_j / E_j$. Equation (5.3) gives the percentage of current income that household j would be willing to pay in order to move from its current location to the immediate vicinity of the market.

Estimates of the θ^h and the β^h are available from regressions presented Tables 3-7. What we need to solve equation (5.3) for c_j are relative ranking weights γ^h. Since we do not have a direct subjective measure of V^h, ranking weights cannot be obtained and a point estimate for C_j cannot be computed. However, equation (5.3) enables us to bracket c_j: by trying various combinations of relative ranking weights γ^h's, we can compute an upper and lower bound for c_j. For instance, if we set $\gamma^{food} = 1$ and other $\gamma^{h \neq food} = 0$, we obtain one possible estimate of c_j as:

$$\begin{aligned} \log(1 - c_j^{food}) &= \frac{\theta^{food}}{\beta^{food}} (\log d^{\min} - \log d_j) \\ c_j^{food} &= 1 - e^{\frac{\theta^{food}}{\beta^{food}} (\log d^{\min} - \log d_j)} \end{aligned} \tag{5.4}$$

Similar c_j^h values can be obtained for other goods.

As is clear from (5.3) and (5.4), computing the welfare cost of isolation ultimately involves dividing the distance coefficient θ^h by the consumption coefficient β^h. Obtaining a consistent estimate of c_j thus depends on getting consistent estimates of θ^h and β^h. So far we have focused primarily on obtaining a consistent estimate of θ^h. We also need to worry about β^h. To this we turn now.

Two potential sources of endogeneity require serious attention: measurement error and unobserved heterogeneity. Because measurement error typically biases coefficient estimates towards zero, its presence would lead to an overestimation of the cost of isolation since we would be dividing by a smaller β^h. It thus important to correct for measurement error in order to obtain an accurate estimate of c_j^h. Unobserved heterogeneity also remains a source of concern. It is indeed conceivable – even likely – that individuals with a grumpy disposition earn less than more cheerful individuals, and hence consume less. Indeed people suffering from depression often antagonize those around them. This is likely to have a negative impact on their income earning potential.

To correct for endogeneity, we need to instrument consumption expenditures so as to correct for measurement error while eliminating from consumption expenditures the possible effect of a grumpy disposition. For this reason, instruments must not include variables that may be affected by someone's grumpiness, such as the size of their family, their current assets, etc. To this effect, we only use instrumental variables that can reasonably be regarded as pre-determined from the individual's perspective – such as their parental background, age, education, and ethnicity. We also need to capture variations in consumption over time. To this effect, we use the mean and standard deviation of rainfall as additional instrument. To the extent that different households are affected by rainfall differently, we interact the rainfall variables with family background. The instrumenting equation is shown in Table 8. Our instruments are jointly significant with a

p-value of 0. We also conduct an overidentification test.[24] In all cases, we cannot reject the null hypothesis that instruments can be excluded from the adequacy regressions. The p-values are 46% and above.

The model presented in Table 6 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. 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 6 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 (5.3) for each V_j^h separately and for a combined measure with equal weight γ^h to all V_j^h. Table 10 reports point estimates of c_j^h estimated at the average and median distance to markets. As it turns out, the two figures are virtually identical. A 95% confidence interval is also reported for c_j^h.[25]

We see that the welfare cost of isolation is very large. On average, if all surveyed households could be brought within 5 minutes walk from the market, one could reduce consumption expenditures on average by 28% while keeping income adequacy unchanged. If we are willing to assume that answers to income adequacy questions capture utility rankings, this is the same as saying that, across our sample, the average willingness to pay for eliminating isolation is 28% of

[24] To implement the overidentification test, we temporarily ignore the selection issue, in which case the standard Sargan test can be used.

[25] The confidence interval is obtained by noting that, for a given value of d_j such as the mean or median, equation (5.4) is nothing but a non-linear function of estimated paramters and hence its confidence interval can be estimated in the usual way. A confidence interval cannot be estimated when coefficients from different regressions are combined in computing (5.3), as when we set $\gamma^h = 1/6$ for all h.

consumption expenditures. Figures are higher when we base our estimate of c_j on consumption adequacy results. If we were to bring everyone within 5 minutes of the market, estimated coefficients predict that we could keep food consumption adequacy assessments unchanged and reduce consumption expenditures on average by 42%. Equivalent figures for housing and schooling are even higher at 65% and 61%, respectively. The highest willingness to pay to reduce isolation is in terms of access to health care. This is probably because households living in isolated wards find it difficult to obtain health care in case of medical emergencies. This is also a reflection of the fact that, in Nepal, income is not a major determinant of access to health care; geographical isolation is.

Comparing the welfare cost of isolation across categories confirms our earlier suspicion that respondents factor the utility gain from variety in their reported adequacy of food and clothing consumption. If also suggests that access to public services such as schooling and, more importantly, health care represent a large proportion of the subjective welfare loss generated by isolation. These findings imply that welfare assessments based on geographical poverty maps massively underestimate the subjective welfare cost of isolation since these maps typically focus solely on monetary income and consumption.

6. Conclusion

Using 1995/96 household survey data from Nepal, we have estimated the subjective welfare cost associated with isolation. 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 so recent and so much of the country remains inaccessible.[26]

[26] The first road to reach Katmandu was built in 1929.

We find the subjective cost of isolation to be significant and very large. The results presented here also provide some evidence on the relative importance of physical distance, consumption variety, and access to public services in the subjective cost of isolation. It appears that, of the three, access to public services is the largest contributing factor. Consumption variety is also an important factor.

Central to any study of the relative attractiveness of various geographical locations is the question of the spatial mobility of individuals. The evidence presented here indicates that, during the study period, a large proportion of Nepalese households could not move freely and costlessly. This hardly surprising given the mountainous nature of the country and the insufficient availability of roads. Nepalese people do not live in isolated mountain hamlets by choice but because they were born there. Our results predict that, given time and opportunity, many of them will 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 traditional explanation for rural-urban migrations in developing countries is the search for higher incomes and better employment. While this explanation may be relevant in some cases, the evidence provided here suggests another possible explanation, namely that rural dwellers are attracted by the amenities 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).

Brickman, P. & D.T. Campbell. 1971. Hedonic Relativism and Planning the Good Society. In *Adaptation-level Theory: A Symposium.* New York: M.H. Apley (ed.), Academic Press pp. 287–302.

Curtis, Eaton B. & Mukesh Eswaran. 2003. "The Evolution of Preferences and Competition: A Rationalization of Veblen's Theory of Invidious Comparisons." *Canadian Journal of Economics* 36(4):832–59.

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.

Di Tella, Rafael, Robert J. MacCullock & Andrew J. Oswald. 2001. "Preferences over Inflation and Unemployment: Evidence from Surveys of Happiness." *American Economic Journal* 91:335–41.

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, Marissa Diener & Carol Diener. 1995. "Factors Predicting the Subjective Well-Being of Nations." *Journal of Personality Social Psychology* 69(5):851–64.

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

Diener, Ed & Shigehiro Oishi. 2000. Money and Happiness: Income and Subjective Well-Being Across Nations. In *Culture and Subjective Well-Being*. Cambridge Mass.: Ed Diener and Eunkook M. Suh (eds.), MIT Press pp. 185–218.

Duesenberry, James. 1949. *Income, Savings and the Theory of Consumer Behavior*. Cambridge Mass.: Harvard University Press.

Easterlin, Richard A. 1974. Does Economic Growth Improve the Human Lot? Some Empirical Evidence. In *Nations and Households in Economic Growth: Essays in Honor of Moses Abramowitz*. New York: Paul A. David and Melvin Reder (eds.), Academid Press pp. 89–125.

Easterlin, Richard A. 1995. "Will Raising the Incomes of All Increase the Happiness of All?" *Journal of Economic Behavior and Organization* 27(1):35–48.

Easterlin, Richard A. 2001. "Income and Happiness: Towards a Unified Theory." *Economic Journal* 111:465–84.

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

Fafchamps, Marcel. 2003. *Rural Poverty, Risk and Development*. Cheltenham (UK): Edward Elgar Publishing.

Fafchamps, Marcel & Agnes R. Quisumbing. 1999. "Human Capital, Productivity, and Labor Allocation in Rural Pakistan." *Journal of Human Resources* 34(2):369–406.

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 & Bart Minten. 2004. "Crime, Transitory Poverty, Isolation: Evidence from Madagascar." (mimeograph).

Fafchamps, Marcel & Christine Moser. 2003. "Crime, Isolation, and Law Enforcement." *Journal of African Economies* 12(4):625–71.

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

Fafchamps, Marcel & Forhad Shilpi. 2004. "Subjective Well-being, Isolation, and Rivalry." (mimeograph).

Fafchamps, Marcel & Forhad Shilpi. 2005. "Cities and Specialization: Evidence from South Asia." *Economic Journal* . (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. 2000. "Happiness, Economy, and Institutions." *Economic Journal* 110:918–38.

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

Fujita, Masahisa, Paul Krugman & Anthony J. Venables. 1999. *The Spatial Economy: Cities, Regions, and International Trade*. Cambridge and London: MIT Press.

Gugler, Josef. 1991. "Life in a Dual System Revisited: Urban-Rural Ties in Enugu, Nigeria, 1961-87." *World Development* 19 (5):399–409.

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

Inglehart, R. & H-D. Klingemann. 2000. Genes, Culture, Democracy, and Happiness. In *Subjective Well-Being Across Cultures.* Cambridge Mass.: E. Diener and E. Suh (eds.), MIT Press.

Inglehart, Ronald F. 1990. *Culture Shift in Advanced Industrial Society.* Princeton: Princeton University Press.

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.

Kingdon, Geeta G. & John Knight. 2004. *Community, Comparisons and Subjective Well-being in a Divided Society.* Oxford: CSAE, Oxford University. WPS 2004-21.

Kingdon, Geeta & John Knight. 2003. "Well-Being Poverty vs Income Poverty and Capabilities Poverty." (mimeograph).

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.

Lokshin, Michael, Nithin Umapathi & Stefano Paternostro. 2003. "Robustness of Subjective Welfare Analysis in a Poor Developing Country: Madagascar 2001." (mimeograph).

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 & Michael Lokshin. 2001. "Identifying Welfare Effects from Subjective Questions." *Economica* 68:335–57.

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.

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.

Venables, Anthony J. & Nuno Limao. 2002. "Geographical Disadvantage: a Hecksher-Ohlin-von Thunen Model of International Specialization." *Journal of International Economics* 58(2):239–63.

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

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

A. Household variables		**Mean**	**Median**	**St.dev.**	**Minimum**	**Maximum**
Travel time to local markets	Hours	2.2	1.1	3.4	0.0	40.0
Total annual consumption	US$	862	563	1015	29	19940
Total value of assets	US$	9910	2445	29854	0	714789
Age of head	Years	44.8	43.0	14.4	11.0	92.0
Education of Head	Years	3.6	0.0	5.0	0.0	17.0
Household Size	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
Days lost due to sickness during last year	Days	2.3	0	6.57	0	70
Travel time to nearest school	Hours	0.37	0.25	0.88	0	25
Travel time to nearest health facility	Hours	1.08	0.5	1.78	0	30
Female heads household	Female=1	13.6%				
Not living in birth ward		19.5%				
One or more household members unemployed		18.0%				
Household hires permanent or casual labor		80.1%				
Parental background						
Inherited land	Hectares	0.81	0.36	1.69	0.00	32.05
Educaton of father of household head	Years	0.9	0.0	2.5	0.0	16.0
Father of head employed in non-farm job	Yes=1	17.0%				
Ethnicity Dummy						
Chhetry (upper caste Hindu)	Yes=1	19.7%				
Brahmin (upper caste Hindu)	Yes=1	16.1%				
Newar (tribal; with own caste system)	Yes=1	9.4%				
Tharu (tribal)	Yes=1	5.8%				
Magar (tribal)	Yes=1	5.0%				
Tamang (tribal)	Yes=1	4.4%				
Gurung (tribal)	Yes=1	4.0%				
Muslim	Yes=1	3.9%				
Yadav/Ah (middle caste Hindu)	Yes=1	3.1%				
Damai (lower caste Hindu)	Yes=1	1.8%				
Kami (lower caste hindu)	Yes=1	4.8%				
Sarki (lower caste Hindu)	Yes=1	1.5%				
Rai (part of Kirat religion)	Yes=1	1.7%				
Limbu (part of Kirat religion)	Yes=1	1.9%				
Other	Yes=1	16.9%				
B. Ward variables						
Urban population within 2 hours travel time of ward	Thousands	53.2	22.9	74.0	0.0	247.5
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
Ward inequality in total consumption	Gini coef.	0.290	0.282	0.083	0.083	0.609
Ward inequality in per capita consumption	Gini coef.	0.257	0.246	0.082	0.091	0.509
Variety						
Food products	Index	0.12	0.09	0.07	0.03	0.49
Non-food products	Index	0.14	0.14	0.03	0.09	0.42
Durable goods	Index	0.44	0.44	0.16	0.15	0.85
Ecological belt dummies						
Mountain	Yes=1	12.1%				
Hills	Yes=1	51.6%				
Terai	Yes=1	36.3%				
Regional dummies						
Eastern	Yes=1	21.1%				
Central	Yes=1	39.2%				
Western	Yes=1	18.6%				
Midwest	Yes=1	10.6%				
Farwest	Yes=1	10.4%				
Maoist Insurgency						
Most affected		2.5%				
Affected		9.9%				
Little affected		20.4%				
Not affected		67.2%				
Rainfall						
Average rainfall	Millimeters	1702	1459	612	1039	3431
Standard deviation of rainfall in ward	Millimeters	411	366	197	176	903
Number of observations		3337				

Table 3. Regressing income and consumption adequacy on isolation and household characteristics

	Dependent variables					
	Total	Consumption of:			Access to:	
Isolation	**income**	**Food**	**Clothing**	**Housing**	**Schooling**	**Health care**
Travel time to local markets (log)	-0.07	-0.20	-0.14	-0.20	-0.23	-0.28
	(2.16)*	(6.29)**	(4.75)**	(6.51)**	(6.92)**	(9.22)**
Proximity to urban centers	2.28	5.29	3.06	0.76	2.05	3.63
	(4.56)**	(9.95)**	(5.95)**	-1.56	(3.77)**	(6.86)**
Household consumption and controls						
Consumption expenditures (log)	0.28	0.51	0.45	0.17	0.30	0.07
	(5.35)**	(9.43)**	(8.52)**	(3.32)**	(5.20)**	-1.43
Value of assets (log)	0.13	0.16	0.12	0.11	0.11	0.10
	(8.05)**	(10.48)**	(8.61)**	(7.83)**	(6.38)**	(7.10)**
Household size (log)	-0.17	-0.36	-0.30	-0.04	-0.29	-0.08
	(2.22)*	(4.66)**	(3.95)**	-0.52	(3.32)**	-1.01
Share of adult females	-0.08	-0.02	-0.03	-0.15	0.17	-0.00
	-0.33	-0.10	-0.15	-0.68	-0.56	-0.02
Share of children under 6	-0.42	-0.44	-0.11	-0.64	-0.10	-0.53
	-1.63	-1.69	-0.42	(2.55)*	-0.34	(2.01)*
Share of youths aged 6 to 20	-0.60	-0.35	-0.40	-0.62	-0.01	-0.33
	(2.96)**	-1.68	-1.92	(3.14)**	-0.04	-1.56
Share of elderly	0.15	-0.42	0.17	-0.09	0.25	-0.14
	-0.57	-1.62	-0.66	-0.36	-0.73	-0.55
Age of household head	-0.00	-0.02	-0.01	-0.02	-0.04	-0.02
	-0.19	-1.95	-0.68	-1.86	(3.37)**	-1.77
Age of household head squared	0.00	0.00	0.00	0.00	0.00	0.00
	-0.13	-1.95	-0.66	-1.76	(3.20)**	-1.54
Female head dummy	-0.03	-0.03	0.04	0.05	-0.03	0.04
	-0.31	-0.34	-0.46	-0.62	-0.30	-0.44
Prices						
Median wage rate in ward (log)	0.10	-0.19	0.05	-0.01	0.01	0.11
	(2.06)*	(3.96)**	-1.13	-0.27	-0.09	(2.23)*
Median rice price in ward (log)	-0.10	-0.24	0.04	0.29	0.01	0.16
	-1.03	(2.42)*	-0.38	(3.05)**	-0.05	-1.63
Regional dummies (Terai and Central region are omitted categories)						
Mountain	-0.31	-0.18	-0.28	-0.11	-0.18	0.26
	(2.97)**	-1.89	(2.87)**	-1.12	-1.64	(2.67)**
Hills	-0.11	-0.21	-0.28	-0.13	-0.24	0.12
	-1.70	(3.20)**	(4.36)**	(2.01)*	(3.41)**	-1.94
Eastern	0.33	0.41	0.27	0.12	0.55	0.27
	(4.39)**	(5.61)**	(3.74)**	-1.69	(6.69)**	(3.86)**
Western	-0.26	0.11	-0.18	-0.22	-0.15	-0.19
	(3.30)**	-1.48	(2.44)*	(3.03)**	-1.79	(2.52)*
Midwest	0.31	0.66	0.35	0.00	0.61	0.33
	(3.24)**	(6.99)**	(3.83)**	-0.01	(6.07)**	(3.70)**
Farwest	0.39	0.38	0.05	-0.16	0.05	-0.09
	(3.97)**	(4.01)**	-0.48	-1.65	-0.46	-0.95
Intercept	-4.41	-4.76	-5.78	-3.23	-3.20	-2.39
	(7.74)**	(8.15)**	(10.04)**	(5.89)**	(4.95)**	(4.23)**
Number of observations	3237	3247	3245	3242	2608	3227
Log likelihood	-1740.12	-1793.17	-1819.62	-1927.30	-1507.62	-1821.20

Absolute value of t statistics in parentheses: * significant at 5%; ** significant at 1%

Table 4. Regressions of income and consumption adequacy controlling for migration selection

	Dependent variables					
	Total	Consumption of:			Access to:	
Isolation	**income**	**Food**	**Clothing**	**Housing**	**Schooling**	**Health care**
Travel time to local markets (log)	-0.07	-0.17	-0.11	-0.19	-0.24	-0.29
	(2.02)*	(5.03)**	(3.19)**	(5.17)**	(6.01)**	(7.82)**
Proximity to urban centers	1.70	3.65	2.00	0.30	1.22	3.22
	(2.71)**	(5.83)**	(3.20)**	-0.50	-1.84	(4.87)**
Household consumption and controls						
Consumption expenditures (log)	0.18	0.41	0.35	0.09	0.19	-0.01
	(2.84)**	(6.72)**	(5.36)**	-1.52	(2.67)**	-0.10
Value of assets (log)	0.20	0.21	0.18	0.16	0.18	0.13
	(8.70)**	(9.06)**	(8.77)**	(7.93)**	(7.57)**	(6.82)**
Household size (log)	-0.10	-0.32	-0.17	0.04	-0.10	0.02
	-1.16	(3.65)**	-1.89	-0.44	-0.99	-0.23
Share of adult females	-0.25	-0.17	-0.06	-0.08	0.63	-0.02
	-0.87	-0.60	-0.20	-0.29	-1.69	-0.08
Share of children under 6	-0.74	-0.50	-0.21	-0.73	0.04	-0.71
	(2.40)*	-1.67	-0.70	(2.46)*	-0.11	(2.32)*
Share of youths aged 6 to 20	-0.67	-0.34	-0.33	-0.43	0.22	-0.33
	(2.73)**	-1.38	-1.36	-1.81	-0.75	-1.33
Share of elderly	0.09	-0.33	0.12	-0.14	0.34	-0.10
	-0.28	-1.05	-0.38	-0.43	-0.82	-0.32
Age of household head	-0.01	-0.02	-0.00	-0.02	-0.04	-0.02
	-0.49	-1.26	-0.33	-1.75	(2.59)**	-1.75
Age of household head squared	0.00	0.00	0.00	0.00	0.00	0.00
	-0.25	-1.31	-0.33	-1.69	(2.43)*	-1.59
Female head dummy	0.08	0.10	0.11	0.05	-0.05	0.00
	-0.71	-0.94	-1.02	-0.51	-0.44	-0.04
Prices						
Median wage rate in ward (log)	0.01	-0.17	-0.00	-0.05	-0.08	0.07
	-0.19	(2.93)**	-0.06	-0.81	-1.24	-1.15
Median rice price in ward (log)	-0.16	-0.20	0.00	0.31	-0.18	0.13
	-1.40	-1.85	-0.03	(2.85)**	-1.47	-1.19
Regional dummies (Terai and Central region are omitted categories)						
Mountain	-0.42	-0.38	-0.41	-0.17	-0.12	0.23
	(3.69)**	(3.62)**	(3.81)**	-1.59	-1.02	(2.12)*
Hills	-0.26	-0.36	-0.39	-0.19	-0.19	0.11
	(3.30)**	(4.68)**	(5.22)**	(2.53)*	(2.26)*	-1.40
Eastern	0.32	0.30	0.25	0.02	0.46	0.20
	(3.75)**	(3.75)**	(3.14)**	-0.22	(4.98)**	(2.41)*
Western	-0.26	0.06	-0.19	-0.27	-0.25	-0.19
	(2.90)**	-0.66	(2.30)*	(3.17)**	(2.61)**	(2.19)*
Midwest	0.29	0.48	0.28	-0.13	0.52	0.31
	(2.62)**	(4.51)**	(2.59)**	-1.20	(4.34)**	(2.90)**
Farwest	0.45	0.33	0.09	-0.17	-0.01	-0.09
	(3.94)**	(3.13)**	-0.84	-1.56	-0.07	-0.80
Intercept	-3.73	-4.44	-5.43	-3.12	-2.83	-1.83
	(5.42)**	(6.48)**	(7.78)**	(4.74)**	(3.70)**	(2.74)**
Number of observations	2406	2412	2412	2408	2509	2396
Log likelihood	-2591.11	-2693.58	-2706.42	-2732.40	-2341.49	-2706.57
Likelihood ratio test of independence between the selection and satisfaction equations						
Chi2(1) test-statistic	2.10	4.71	3.20	2.41	2.43	0.36
P-value	0.15	0.03	0.07	0.12	0.12	0.55

Absolute value of t statistics in parentheses: * significant at 5%; ** significant at 1%

Table 5. Non-migration selection regressions

The dependent variable is 1 if the head of household resides in birth ward, 0 otherwise.

(Each selection regression corresponds to an adequacy regression since they are estimated together)

	Selection regression corresponding to adequacy regression for:					
	Total	**Consumption of:**			**Access to:**	
	income	**Food**	**Clothing**	**Housing**	**Schooling**	**Health care**
Age of household head	-0.00	-0.00	-0.00	0.00	0.01	-0.00
	-0.06	-0.09	-0.15	-0.02	-1.00	-0.16
Age of household head squared	0.00	0.00	0.00	0.00	0.00	0.00
	-0.46	-0.39	-0.36	-0.49	-1.44	-0.35
Log(education of household head)	-0.12	-0.10	-0.13	-0.13	-0.09	-0.12
	(4.05)**	(3.25)**	(4.26)**	(4.13)**	(2.78)**	(3.95)**
Father employed in non-farm job	-0.03	0.00	-0.05	-0.03	-0.00	-0.03
	-0.43	-0.06	-0.61	-0.41	-0.03	-0.37
Education of father of head (log)	-0.28	-0.27	-0.28	-0.29	-0.30	-0.28
	(6.23)**	(6.11)**	(6.09)**	(6.38)**	(6.31)**	(6.10)**
Inherited land (log)	0.72	0.77	0.70	0.72	0.81	0.73
	(9.30)**	(10.32)**	(8.81)**	(9.39)**	(9.96)**	(9.51)**
Ethnicity dummies	yes	yes	yes	yes	yes	yes
Intercept	1.17	1.15	1.19	1.17	0.54	1.19
	(4.34)**	(4.24)**	(4.40)**	(4.31)**	-1.84	(4.37)**
Number of observations	2973	2979	2979	2975	2509	2963

Absolute value of t statistics in parentheses: * significant at 5%; ** significant at 1%

Table 6. Adequacy regressions with additional controls
(All regressions control for self-selection but self-selection regressions not shown to save space)

	Dependent variable is adequacy of:					
	Total	Consumption of:			Access to:	
Isolation	**income**	**Food**	**Clothing**	**Housing**	**Schooling**	**Health care**
Travel time to local markets (log)	-0.06	-0.12	-0.11	-0.19	-0.25	-0.28
	-1.53	(3.38)**	(2.85)**	(4.79)**	(5.68)**	(6.94)**
Proximity to urban centers	2.55	3.91	2.96	1.21	2.01	3.66
	(3.48)**	(5.72)**	(4.13)**	-1.71	(2.48)*	(5.01)**
Household consumption and controls						
Consumption expenditures (log)	0.22	0.38	0.38	0.13	0.22	0.05
	(3.16)**	(6.08)**	(5.70)**	(1.98)*	(2.91)**	-0.82
Value of assets (log)	0.17	0.16	0.14	0.13	0.16	0.11
	(6.05)**	(6.50)**	(6.14)**	(5.99)**	(5.65)**	(4.99)**
Household size (log)	-0.05	-0.24	-0.13	0.08	-0.09	0.03
	-0.54	(2.80)**	-1.45	-0.82	-0.87	-0.26
Share of adult females	-0.39	-0.19	-0.09	-0.10	0.64	-0.01
	-1.27	-0.69	-0.30	-0.35	-1.63	-0.03
Share of children under 6	-0.86	-0.44	-0.20	-0.72	0.16	-0.55
	(2.63)**	-1.50	-0.63	(2.31)*	-0.41	-1.74
Share of youths aged 6 to 20	-0.82	-0.35	-0.34	-0.46	0.25	-0.31
	(3.13)**	-1.48	-1.34	-1.84	-0.78	-1.20
Share of elderly	0.03	-0.51	0.00	-0.31	0.28	-0.19
	-0.10	-1.66	-0.01	-0.93	-0.64	-0.58
Age of household head	0.00	-0.01	0.01	-0.02	-0.03	-0.01
	-0.03	-0.65	-0.38	-1.17	(2.08)*	-0.75
Age of household head squared	0.00	0.00	0.00	0.00	0.00	0.00
	-0.04	-1.19	-0.02	-1.39	(2.27)*	-0.94
Female head dummy	0.18	0.30	0.23	0.15	0.06	0.11
	-1.50	(2.85)**	(2.00)*	-1.36	-0.47	-0.95
Prices						
Median wage rate in ward (log)	0.00	-0.18	-0.01	0.01	-0.13	0.08
	-0.01	(3.02)**	-0.11	-0.19	-1.80	-1.19
Median rice price in ward (log)	-0.22	-0.27	-0.00	0.22	-0.29	0.02
	-1.72	(2.41)*	-0.02	-1.78	(2.06)*	-0.18
Additional controls						
Travel time to nearest school (log)	-0.05	-0.25	-0.04	0.06	-0.27	-0.09
	-0.33	(2.03)*	-0.33	-0.45	-1.83	-0.68
Travel time to nearest health facility (log)	-0.09	-0.15	-0.04	-0.05	-0.01	-0.05
	-1.05	(2.24)*	-0.48	-0.66	-0.07	-0.63
Variety index for food	-0.33	0.49	-0.13	0.01	1.10	-0.20
	-0.57	-0.93	-0.23	-0.02	-1.58	-0.36
Variety index for non-food	1.84	2.61	0.62	2.90	0.87	2.86
	(2.04)*	(3.03)**	-0.69	(3.23)**	-0.87	(3.15)**
Variety index for durables	0.22	-0.05	0.53	0.26	0.54	0.17
	-0.76	-0.18	-1.93	-0.95	-1.72	-0.60
Education of household head (log)	0.13	0.22	0.18	0.13	0.17	0.14
	(3.52)**	(6.85)**	(5.16)**	(3.31)**	(4.23)**	(3.83)**
One or more household members unemployec	-0.43	-0.23	-0.34	-0.35	-0.22	-0.27
	(5.04)**	(3.19)**	(4.42)**	(4.50)**	(2.51)*	(3.44)**
Days lost to illness over preceding year	-0.02	-0.01	-0.01	-0.01	-0.01	-0.02
	(3.13)**	(2.51)*	-1.44	-1.87	-1.85	(4.57)**
Ward inequality in per capita consumption	-0.23	0.26	0.56	0.48	1.36	0.20
	-0.59	-0.74	-1.49	-1.28	(3.11)**	-0.51
Average rainfall in ward	0.00	0.00	0.00	0.00	0.00	0.00
	-1.02	-1.69	-1.91	(3.16)**	(4.37)**	(4.01)**
Standard deviation of rainfall in ward	0.00	0.00	0.00	0.00	0.00	0.00
	-0.24	(3.17)**	-0.93	(4.07)**	(2.28)*	(2.40)*
Districts most affected by Maoist insurgency	-0.01	0.39	0.15	0.12	0.26	0.25
	-0.03	(1.97)*	-0.70	-0.54	-1.10	-1.20
Districts affected by Maoist insurgency	0.47	0.10	0.28	0.50	0.34	0.51
	(3.79)**	-0.92	(2.39)*	(4.21)**	(2.64)**	(4.30)**
Districts little affected by Maoist insurgency	-0.02	-0.06	-0.15	-0.18	-0.21	-0.14
	-0.22	-0.70	-1.61	(1.97)*	(2.08)*	-1.48
Household hires permanent or casual labor	-0.11	0.14	-0.05	0.17	-0.20	-0.17
	-1.11	-1.61	-0.50	-1.79	-1.72	-1.81
Regional dummies			included but not shown			
Ethnicity dummies			included but not shown			
Intercept	-3.53	-4.31	-5.65	-4.11	-2.86	-2.28
	(4.10)**	(5.40)**	(6.67)**	(4.95)**	(2.93)**	(2.71)**
Log likelihood	-2531.04	-2619.47	-2652.04	-2649.52	-2268.05	-2607.88
Likelihood ratio test of independence between the selection and satisfaction equations						
Chi2(1) test-statistic	0.16	12.43	0.36	0.52	0.06	1.17
P-value	0.69	0.00	0.55	0.47	0.80	0.28

Absolute value of t statistics in parentheses: * significant at 5%; ** significant at 1%

Table 7. Adequacy regressions with average ward distance to markets

	Dependent variable is adequacy of:					
	Total income	Consumption of:			Access to:	
		Food	Clothing	Housing	Schooling	Health care
Average travel time to markets in ward	-0.11	-0.21	-0.17	-0.21	-0.34	-0.38
	(2.25)*	(4.94)**	(3.72)**	(4.69)**	(6.82)**	(8.36)**
	(Other regressors as in Table 6 – not shown here to save space)					

Absolute value of t statistics in parentheses: * significant at 5%; ** significant at 1%

Table 8. Instrumenting regression for consumption expenditures

Pre-determined variables	
Age of household head	0.04
	(7.49)**
Age of household head squared	0.00
	(4.72)**
Log(education of household head)	0.27
	(20.84)**
Instruments	
Education of father of head (log)	0.22
	(4.14)**
Father employed in non-farm job	0.03
	-0.35
Rainfall x father in non-farm job	0.00
	(3.61)**
St.dev. of rainfall x father in non-farm job	-0.00
	(3.04)**
Rainfall x father's education	0.00
	-0.05
St.dev. of rainfall x father's education	0.00
	-1.91
Ethnicity dummies	included but not shown
Intercept	8.93
	(77.83)**
Number of observations	3021
R-squared	0.30
Joint F-test of the instruments	
Chi2 test-statistic	19.91
P-value	0.00

Absolute value of t statistics in parentheses: * significant at 5%; ** significant at 1%

Table 9. Satisfaction regressions with instrumented consumption expenditures

	Dependent variables					
	Total	Consumption of:			Access to:	
Isolation	**income**	**Food**	**Clothing**	**Housing**	**Schooling**	**Health care**
Travel time to local markets (log)	-0.06	-0.12	-0.10	-0.19	-0.24	-0.27
	-1.49	(3.30)**	(2.76)**	(4.82)**	(5.45)**	(6.76)**
Proximity to urban centers	2.54	3.86	2.94	1.26	2.06	3.61
	(3.54)**	(5.79)**	(4.19)**	-1.78	(2.62)**	(5.00)**
Household consumption and controls						
Consumption expenditures (log)	0.88	1.04	1.13	0.87	1.23	0.45
	(3.42)**	(4.28)**	(4.38)**	(3.19)**	(4.32)**	-1.63
Residuals from instrumenting regression	-0.68	-0.68	-0.78	-0.76	-1.04	-0.40
	(2.66)**	(2.82)**	(3.01)**	(2.79)**	(3.68)**	-1.47
Value of assets (log)	0.16	0.15	0.13	0.13	0.14	0.11
	(5.45)**	(6.32)**	(5.63)**	(5.38)**	(4.85)**	(4.75)**
Household size (log)	-0.05	-0.22	-0.12	0.08	-0.09	0.03
	-0.50	(2.64)**	-1.37	-0.82	-0.85	-0.28
Share of adult females	-0.37	-0.17	-0.07	-0.09	0.69	-0.01
	-1.24	-0.64	-0.24	-0.32	-1.80	-0.02
Share of children under 6	-0.83	-0.43	-0.19	-0.71	0.20	-0.54
	(2.58)**	-1.49	-0.62	(2.25)*	-0.53	-1.73
Share of youths aged 6 to 20	-0.79	-0.34	-0.33	-0.46	0.29	-0.30
	(3.08)**	-1.47	-1.32	-1.81	-0.95	-1.19
Share of elderly	0.08	-0.45	0.05	-0.26	0.40	-0.17
	-0.23	-1.51	-0.17	-0.79	-0.93	-0.52
Age of household head	-0.02	-0.03	-0.02	-0.04	-0.07	-0.02
	-1.53	(2.14)*	-1.42	(2.60)**	(3.86)**	-1.48
Age of household head squared	0.00	0.00	0.00	0.00	0.00	0.00
	-1.08	(2.19)*	-1.17	(2.43)*	(3.61)**	-1.46
Female head dummy	0.16	0.28	0.21	0.15	0.04	0.10
	-1.43	(2.77)**	-1.88	-1.28	-0.31	-0.91
Prices						
Median wage rate in ward (log)	-0.00	-0.18	-0.01	0.01	-0.13	0.07
	-0.05	(3.09)**	-0.17	-0.16	-1.88	-1.18
Median rice price in ward (log)	-0.20	-0.26	0.01	0.23	-0.26	0.03
	-1.62	(2.32)*	-0.07	-1.89	-1.88	-0.23
Additional controls						
Travel time to nearest school (log)	-0.04	-0.24	-0.04	0.06	-0.25	-0.09
	-0.31	(2.05)*	-0.32	-0.47	-1.81	-0.67
Travel time to nearest health facility (log)	-0.09	-0.16	-0.05	-0.06	-0.02	-0.05
	-1.19	(2.38)*	-0.65	-0.79	-0.22	-0.70
Variety index for food	-0.31	0.50	-0.09	0.03	1.09	-0.17
	-0.54	-0.98	-0.16	-0.06	-1.62	-0.31
Variety index for non-food	1.76	2.57	0.56	2.87	0.82	2.82
	(2.01)*	(3.05)**	-0.64	(3.20)**	-0.84	(3.15)**
Variety index for durables	0.21	-0.04	0.51	0.24	0.50	0.16
	-0.75	-0.15	-1.93	-0.88	-1.64	-0.61
Education of household head (log)	-0.06	0.01	-0.04	-0.09	-0.14	0.02
	-0.76	-0.15	-0.51	-1.01	-1.52	-0.22
One or more household members unemployed	-0.41	-0.22	-0.33	-0.35	-0.20	-0.26
	(4.93)**	(3.11)**	(4.34)**	(4.47)**	(2.43)*	(3.39)**
Days lost to illness over preceding year	-0.02	-0.01	-0.01	-0.01	-0.01	-0.02
	(3.03)**	(2.41)*	-1.34	-1.75	-1.72	(4.46)**
Ward inequality in per capita consumption	-0.20	0.25	0.56	0.50	1.31	0.20
	-0.52	-0.71	-1.52	-1.33	(3.11)**	-0.52
Average rainfall in ward	0.00	0.00	0.00	0.00	0.00	0.00
	-1.26	(2.03)*	(2.24)*	(3.44)**	(4.70)**	(4.14)**
Standard deviation of rainfall in ward	0.00	0.00	0.00	0.00	0.00	0.00
	-0.28	(3.71)**	-1.56	(4.62)**	(3.01)**	(2.68)**
Districts most affected by Maoist insurgency	0.01	0.40	0.17	0.14	0.29	0.26
	-0.06	(2.05)*	-0.82	-0.64	-1.27	-1.25
Districts affected by Maoist insurgency	0.45	0.09	0.27	0.49	0.31	0.50
	(3.64)**	-0.86	(2.30)*	(4.10)**	(2.45)*	(4.21)**
Districts little affected by Maoist insurgency	-0.02	-0.06	-0.14	-0.18	-0.19	-0.13
	-0.17	-0.69	-1.57	-1.92	(2.03)*	-1.45
Household hires permanent or casual labor	-0.09	0.16	-0.03	0.18	-0.18	-0.17
	-0.97	-1.82	-0.33	-1.85	-1.59	-1.76
Regional dummies			included but not shown			
Ethnicity dummies			included but not shown			
Intercept	-9.33	-10.25	-12.30	-10.71	-11.78	-5.79
	(3.97)**	(4.52)**	(5.17)**	(4.30)**	(4.47)**	(2.29)*
Log likelihood	-2527.77	-2615.60	-2647.76	-2645.86	-2261.80	-2606.83
Likelihood ratio test of independence between the selection and satisfaction equations						
Chi2(1) test-statistic	2.23	17.54	2.90	0.40	2.81	2.42
P-value	0.14	0.00	0.09	0.53	0.09	0.12
Overidentification test for consumption expenditures						
Sargan statistic	2.02	3.67	4.05	4.63	4.44	3.49
P-value	0.85	0.60	0.54	0.46	0.49	0.62

Absolute value of t statistics in parentheses: * significant at 5%; ** significant at 1%

Table 10: Compensating Variation evaluated at the mean and median of distance to markets
Expressed in % ot total consumption expenditures

	Evaluated at Mean Distance	(95% confidence interval)		Evaluated at Median Distance	(95% confidence interval)	
Total income	27.8%	-6.8%	62.4%	27.9%	-6.8%	62.6%
Food consumption	41.8%	17.5%	66.2%	42.0%	17.6%	66.4%
Clothing	36.0%	11.3%	60.7%	36.1%	11.3%	60.9%
Housing	64.9%	37.1%	92.8%	65.1%	37.2%	93.0%
School	60.5%	38.2%	82.7%	60.6%	38.4%	82.9%
Health	94.5%	74.1%	114.9%	94.5%	74.3%	114.8%
Equal weights for all 6	54.3%			54.4%		

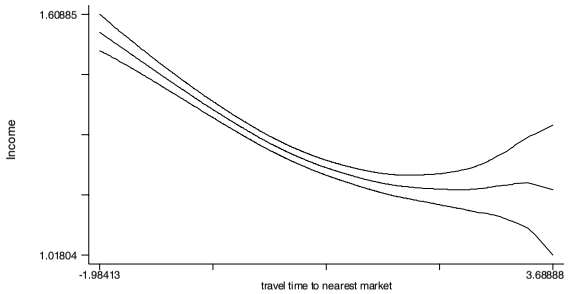

Figure 1a. Total income

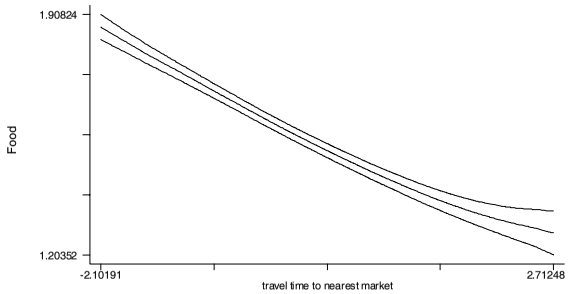

Figure 1b. Food consumption

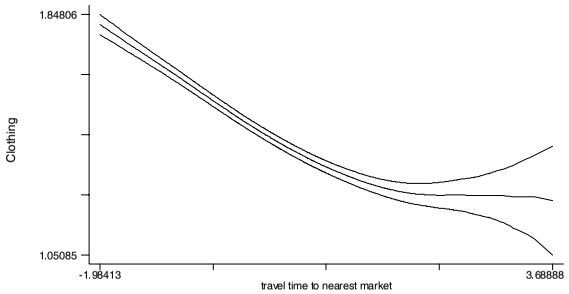

Figure 1c. Clothing consumption

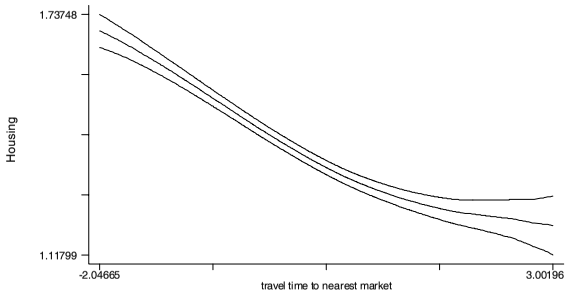

Figure 1d. Housing

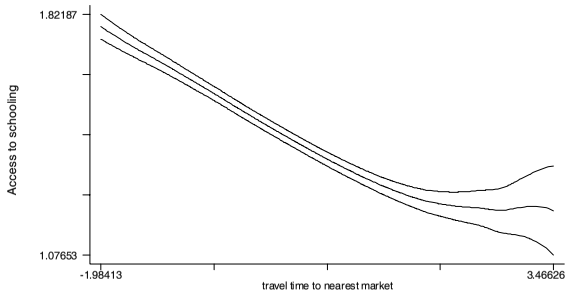

Figure 1e. Access to schooling

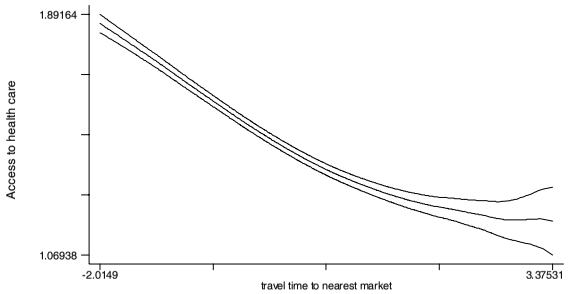

Figure 1f. Access to health care

Figure 1. Non-parametric analysis

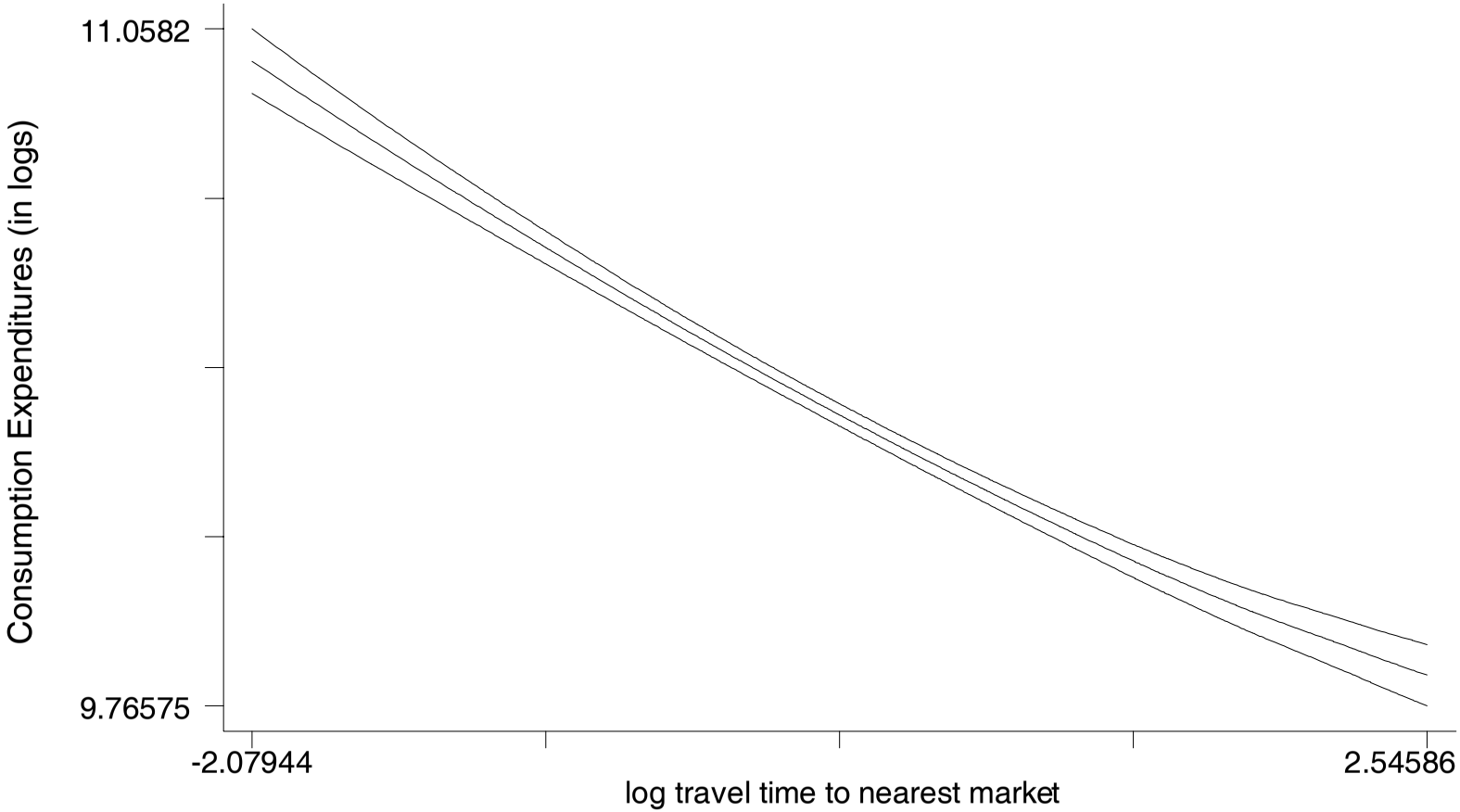

Figure 2. Consumption and isolation