Contemporary Issues and Ideas in Social Sciences
December 2007

Impact of a Highway on the Socio-economic Well-being of Rural Households Living in Proximity

Ramprasad Sengupta[1]
Centre for Economic Studies and Planning
Jawaharlal Nehru University, New Delhi–110 067
Dipankor Coondoo
Economic Research Unit, Indian Statistical Institute, Kolkata–700 035,
and
Bhisma Rout
Asian Development Bank, India Resident Mission, New Delhi–110 021

[1]Corresponding Author; Currently, Visiting Researcher, Research Institute of Economics and Business Administration, (RIEB), Kobe University, Kobe, Japan and Professor of Economics, Jawaharlal Nehru University, New Delhi. Email: rsp0302@gmail.com
The Authors are grateful to the Asian Institute of Transport Development (AITD) and Research Institute for Economics and Business Administration (RIEB), Kobe University for providing support for the preparation and revision of this paper respectively. The idea and methodology of this paper originated from the AITD's project sponsored by the National Highway Authority of India and the World Bank with which the authors were closely associated. The authors would like to acknowledge the encouragement, help and comments received from Mr. K.L. Thapar, Chairman of AITD.

Impact of a Highway on the Socio-economic Well-being of Rural Households Living in Proximity

Abstract

This paper investigates empirically the impact of a highway development on the poverty status and socio-economic conditions of well-being of the people living in its proximity, through greater connectivity and transport mobility. It points out the connection between the road development and the individual welfare maximising choices with the help of a microeconomic rural household model. It shows how the value of the various household level well-being indicators tend to decline as the approach distance of a household from the highway increases and tends to disappear beyond a threshold distance using nonparametric regression models with the baseline survey cross section data pertaining to the project of widening the National Highway 2 of India. The paper further uses propensity score matching cum differencing method to estimate the partial effect of the highway on those welfare indicators and illustratively show how the highway has significant impact on the rural economies in India through which it passes.

1. Introduction

The quality of life and socio-economic conditions of living of people may get significantly affected by the speed and ease with which they can move and carry their goods. In this sense transport and economic development are interdependent and their relationship is both complex and dynamic. Briefly, transport promotes access to markets, materials and opportunities by facilitating movements of persons and goods and improves earning and thereby level of living. This in turn enhances the demand for transport. This two-way interaction works through a host of inter-sectoral forward and backward linkage effects and dynamic externalities, tends to relocate industries, services and labour and thus helps to shape the economic geography of a region.

The ultimate aim of developmental activities, including those relating to transport infrastructure, is to promote societal welfare. However, due to the existing pattern of socio-economic structure, geo-political and historical features etc., the benefits of development are often not shared equitably and a variety of distributional inequalities show up at all levels - local, regional, or national. The greater part of the society often receives little or none of the benefits of development. Typically, this segment consists of the poor who mostly live in rural areas of the developing world.

Development of transport facility like road infrastructure, can, however, play a significant role in changing the socio-economic conditions of living of the people of a region through dynamic externalities that such development often generates. It can in fact be an important element of both direct and indirect interventions for poverty reduction and improvement of socio-economic conditions of the people. However, there has been little assessment of the socio-economic impact of an *infrastructural* project like construction/widening of a highway. It is generally thought that the distributional equity should be dealt with through explicit fiscal policies and other kinds of direct state intervention. It is, however, now being increasingly realised that the socio-economic impact analysis with a thrust on distributional issues like poverty reduction should be made to see how important the role of a transport or a ground infrastructural project may be in bringing about the distributional justice (Levy, 1996; Baker, 2000).

Recently India has embarked upon a programme of upgrading the national highway network connecting the four metropolises and the major maritime ports of the country. This involves massive public investment. Side by side, the country also carries a crushing load of poverty, which is more pronounced in the rural areas. According to the latest estimates, more than one-fourth of the Indian rural population lives below the official poverty line (Deaton and Dreze, 2002). It would, therefore, be worthwhile to examine the socio-economic impact of such massive public investment schemes.

In the above context a study was designed by the Asian Institute of Transport and Development with which the present authors were closely associated to examine:

(a) How living in proximity[2] or accessibility to the National Highway 2 (NH2) (which is one of India's oldest national highway) in its rural stretches between Agra and Dhanbad influence the socio-economic well-being indicators including poverty status for such rural population.,

(b) How the widening of the highway (which is currently on-going) would impact on the socio-economic condition of the same population after the project is completed.

The issue of poverty alleviation gains greater importance in this case because most of the area through which this stretch of NH2 passes has high incidence of rural poverty, according to the official poverty estimates. In this

[2]This stretch of the NH2 between Agra in the state of Uttar Pradesh and Dhanbad in the state of Jharkhand which is 995 km long and passes through the states of Uttar Pradesh, Bihar and Jharkhand, is being four laned from a two lane one.

paper an analytical framework and methodology of socio-economic impact evaluation has been proposed that may be used in cases of such infrastructural development projects as development and up-gradation of highway development. This methodology requires household-level information on a relevant set of *impact* or *outcome* variables collected through two surveys - viz., a pre-project baseline survey and a post-project survey- so that the impact may be estimated by comparing the two sets of cross-section survey data using an appropriate methodology. A cross-sectional baseline survey of the socio – economic conditions of the households of the villages in the rural sections of the above mentioned highway project was conducted to provide the pre-project observational values of a large number of outcome variables representing socio – economic well-being condition, approach distance of the household from the high way and a number of travel related variables.

Apart from proposing the analytical framework and methodology for examining (a) and (b), the paper makes the empirical investigation for the issues pertaining to (a) and estimates the effect of proximity to NH2 on the socio-economic condition of the rural population living in its neighbourhood using the baseline cross-sectional survey data only. As the living condition of the people would also be influenced by a host of other factors, the methodology required for assessing the partial impact of proximity to NH2 as distinct from the other influencing factors has required the delineation of influence zone and control zone in respect of the highway impact and the use of propensity score matching cum single difference method of analysis for such estimation as is explained below.

The influence zone in the model of this paper is defined to be the area adjacent to the highway on its both sides where the well-being indicators of households are expected to be influenced by the approach distance from the high way. The control zone, on the other hand, is chosen to be an area lying beyond the influence zone, where the existence of the highway or its development is supposed to have little or no impact on the well-being indicators of households located in it[3]. This influence zone, however, is to be determined empirically. We have used nonparametric regression technique to examine the nature of relationship between each of a set of selected well-being indicators (which we have called outcome variables here) and the approach distance of a household from the highway and to ascertain the limit of such approach distance up to which the proximity to the road or highway has impact on the socio-economic condition of the people.

While such nonparametric regression has been helpful to delineate the influence zone from the rest of the adjoining area to the highway, the ascertaining of the partial effect of proximity to the highway on the outcome variables would require the use of the propensity score matching method for identifying similar households in the influence zone and control zone. This propensity score matching method identifies for each household in the influence zone a corresponding set of households in the control zone which are very similar in terms of socio-economic characteristics other than the well-being indicators considered. We have applied a single difference method to the values of outcome variables of these comparable or matching households

[3]Note that traditionally the programme evaluation literature identifies a treatment group and a control group. Here we have the households of the influence zone form the treatment group and the corresponding matched households of the control zone form the controlgroup , so to say.

of the two zone to get an estimate of impact of the existing highway on the socio-economic conditions through proximity, i.e., part (a) of the empirical investigation as mentioned above. It is, however, a limitation of the use of single difference method on outcome variables that it cannot completely separate the effects of proximity to highway from the confounding effects of other observable or non-observable factors[4].

Our use of single difference method here would implicitly assume that the other influencing factors of socio-economic well-being are correlated with the variables used for propensity score matching excepting the distance from the highway. This would mean that the similarity of households for matching in the two zones also captures similarity with reference to the effects of other influencing factors as well. This is of course a strong assumption. Besides, it is also to be noted that our enquiry of the impact of proximity to the highway has not included the environmental and health related negative external effects, like the incidence of HIV / AIDs or automotive or road dust pollution related adverse health effects.

Our empirical results in this paper are thus confined to the analysis of the impact of proximity to the existing highway NH2 on the poverty and the level of socio-economic well-being of the people living in its proximity, using the pre-project base line survey cross section data. As the work of the four-

[4]This limitation could have been removed if some pre-existing survey data for the highway on the decided outcome variables for the same identified sample households were available to permit the use of a double difference analysis of the outcome variables to measure the separated partial effects of proximity to the highway. For an already existing highway such Double-Difference analysis of outcome variables of similar household cannot be carried out because the pre-existence of the highway data are not available.

laning of the NH2 is still ongoing and the post project survey is yet to take place, this paper does not address the impact assessment of the widening of the NH2 by four laning which would involve the double difference analysis to obtain the separated partial effect of the highway widening only on the outcome variables.

The paper is organised as follows: The conceptual linkage between highway development and socioeconomic well-being of people is briefly discussed in Section 2; Section 3 describes an analytical model of a rural household that provides a theoretical foundation of the empirical exercise; Section 4 discusses about the methodology of analysis and the data set. We present the empirical results and their developmental implications in Section 5 and, finally, Section 6 concludes the paper.

2. Linkage between Highway Development and Socio-economic Living Condition of People: A Conceptual Framework

The socio-economic impact of development of a *ground-fixed* infrastructure of transport like the NH2 or its widening may be of two types. The direct impact is likely to be in the form of (i) an enhancement of the level of spatial connectivity (and the consequent increase of passenger and freight traffic carrying capacity) which may be initially low and uneven due to socio-economic, physiographic and historical reasons and (ii) a reduction of the cost of provision as well as the cost of use of road infrastructure. These would bring about a higher mobility and lower travel cost (both in terms of money and time) of the people having access to the highway developed.

Rise in the number of trips (for all purposes together) per capita per day of a household or reduction in the average travel cost (if any) for such trips would thus be the immediate direct effects of a highway project. Here the travel cost may be interpreted either as the cost of transport operation along the highway or as the cost of travel and freight movement using the highway. In principle, the transport cost should also include users' time cost, which might be more sensitive than the monetary cost of transportation in the case of highway widening.

The indirect impact of a highway development, on the other hand, would work through the dynamic developmental externalities generated through the forward and the backward linkages[5]. An example of this may be a change in the land use pattern in the areas that get greater connectivity due to the highway, since there will be changes in the patterns of settlement, agricultural land use and location of industries, trading and other services and non-farm unorganised sector activities. All these would be reflected in the changes in the pattern of economic activities, income generation, price evolution, employment conditions and ground rent prevailing in the concerned local region. A new land use pattern may in turn induce greater accessibility to job markets, health and educational facilities, etc., attract investment for the development of feeder roads, power distribution networks, telecommunication facilities and other modes of connectivity among other, leading to a greater access of the local people to markets and infrastructural facilities. All these should have a bearing on the level of well-being of the households, although some of them may not themselves necessarily use more of the highway facility created. These would in turn lead to changes in the level of well-being

[5]These are, in fact, the general equilibrium effects of such an infrastructural development that will work out in terms of induced changes in output and prices across the different sectors over time, having definite welfare implications in the socio-economic context.

and human development, through their impact on consumption level, educational attainment, health status, etc. in the local economies consequent to such road development.

The impact of a highway development may also be seen from a different angle, viz., the local impact and the wider regional or national level impact. The local impact is expected to be limited to the immediate neighbourhood of the highway – that is, to the towns and villages lying on both sides of the highway within a (horizontal or approach) distance of, say, 5 to 7 or 10 km. defining the influence zone. The entire regional or national economy lying beyond this neighbourhood should also benefit from the development. Such effect may be called the regional or national level impact. Further, in either case the impact may be of a direct or indirect nature. It may be mentioned in this context that the indirect *general equilibrium* effects on income, output, employment, land rent and land price, poverty, etc. are realised not only in the local economies in the proximity of the highway, but are also transmitted throughout the regional, if not the national economy by way of various linkage effects. However, we confine our analysis here to the measurement and analysis of the impact of the highway on the local economies in the proximity only[6].

[6] As explained later, for the purpose of analysis we have defined as the influence zone of the highway the area lying within 5 km of horizontal distance on either sides of the highway. We have not tried to carry out an assessment of the national level effects because of the difficulties involved in separating the partial impact of the existence/expansion of a highway from those of a host of other factors driving the process of national development. The absence of a suitable methodology of empirical analysis and the consideration of feasibility of the required data collection prompted us to confine ourselves to the measurement of the local level effects alone.

3. An Analytical Household Model

We propose the following rural household model to identify an appropriate set of indicators of socio-economic well-being and their relation with the proximity to the road infrastructure. It is essentially an optimizing choice model of a representative rural household which is endowed with certain resources of agricultural land and capital. The household rationally allocates the given resources among the alternative uses in order to maximise the utility or welfare (u) subject to the full income and the full time budget constraints. The rural household has mixed income comprising rent from the agricultural land and other assets as well as its wage income. The full time at its disposal is allocated among the uses of working on one's own farm, working as wage labour, travel of all kinds, and leisure. The welfare of the household is dependent on the current consumption of goods and services including services of education and health care, leisure and the future consumption made out of current savings. The travel time and the travel cost of a given household would depend on the parameter of its distance from the highway and also the qualitative state of the highway (that is, if it is two laned or four laned, etc.).

The optimal choice values of the variables would thus be determined by the parameters of the various prices, wages rate, interest rate, land productivity, approach distance of the household (say δ) from the highway and the (qualitative) width status (λ) of the highway (i.e., whether double laned or four laned). Any change in the approach distance parameter, *ceteris paribus*, will have its direct effect on the optimal values of the choice variables like outputs and inputs of agricultural production, labour – leisure choice, consumption of commodities and that of services of education and health care. The relationship between the former and the latter can be traced through the parametric variation of the approach distance parameter. As the other parameter λ of the highway would also influence the optimal choices in this model, one can conceptually perceive how the relationship of any of the optimal choices with the approach distance of the household from the highway would shift with the upgradation of the highway (i.e., with the change in the value of λ).

As the optimal welfare level of a household, u^*, is not observable, we have examined the empirical relationship of each of the subset of *observable* decisions/endogenous variables (expected to be closely related to the level of socio-economic well-being of the sample households[7]) with δ - the approach distance of the household from the highway. It may be mentioned that given a value of λ, δ may affect a household's welfare level in two ways. It may directly affect the optimal levels of the decision/endogenous variables. It may also affect some of the other parameters of the model and thereby indirectly affect the optimal levels of the decision variables. For example, the prices, wage and interest rate are likely to be dependent on the conditions of the local village economy and access to opportunities which may vary systematically with δ. The model further envisages that for any given value of δ, any shift in λ or upgradation of the highway would also influence the endogenous choice variables of the household directly as well as indirectly through their influence on the price – parameters in the long run. The set of outcome variables for the road project like highway development would thus be defined over the domain of the endogenous choice variable of the model and also on some of the price-income related parameters. The formal version of the model is

[7]Elsewhere in this paper we have referred to this set of variables as the impact or outcome variables.

presented in Appendix I. While this formal analytical model does not yield any relationship between the outcome variables or socio-economic well being indices and the distance from the highway in an analytically reduced form, it provides ample rationale for testing the hypothesis that the distance of the location of a household from the highway as well as the qualitative status of the highway like width would matter in determining the level of socio-economic well-being of the rural households located in its neighbourhood.

4. The Methodology and the Data

4.1. Nonparametric Regression

The model as presented above points to the kind of variables that needs to be considered as outcome variables and the rationale of examining the nature of relationship between the δ-parameter of approach distance and each of the individual outcome variables which are influenced directly or indirectly by the former. As the theoretical construct of the model as formally presented in the Appendix I is too general to ascertain the analytical form of such formal relationships, we have used nonparametric regression models to estimate them (Härdle, 1990). This regression method is a powerful tool to estimate the relationships in terms of non analytical functional forms so that their exact nature of behaviour can be captured and their gradient of change can be ascertained. A comparison of this gradient of change as given by the non parametric regression curves between the pre-and the post intervention of a highway widening project, can then be analysed with reference to changes in mobility and accessibility to identify the impact of the project as distinct from other developmental initiatives. (Härdle, 1990; Pagan and Aman Ullah, 1999). Kernel Regressions were used to derive the curves of regression of the outcome variables over the variable of distance of the households from the highway and to obtain the limit of the influence zone[8].

A priori, one would expect the effect of NH2 on the level of the impact variable to decline systematically with the distance[9] from the highway and beyond a *threshold* distance level (δ^*) the effect should cease to exist. The estimated threshold distance from the highway has been used to delineate the influence zone and the survey area lying beyond this influence has been treated as the control zone for the purpose of the present impact analysis with the use of the propensity score matching-cum-differencing method. The control zone would have similar agro-climatic and socio-economic characteristics as the influence zone. The graphical presentations of the nonparametric regression models are given in Figures 1 to 8.

The encompassing distance for the influence zone of a highway may, however, also be defined alternatively in terms of its *accessibility*. For example, any place away from the highway from where the highway can be reached in less than 30 minutes by a bicycle or in one hour on foot (that is, a distance of 4-5 km.) is thought to have access to the highway in an Indian-type situation. Thus, the set of all villages lying within an approach distance of 4-5 km. from the NH2 may be defined to be the influence zone of the NH2. It may be noted that the graphs of the estimated nonparametric regression functions of the individual outcome variables mostly suggest a similar threshold distance and thus justify the delineation of the influence zone for the analysis that follows.

[8]See AITD (2003) for the details of the Kernel Regression methods used in the study.

[9]This is known as the gradient of change hypothesis in the literature of transport studies.

4.2. Propensity Score Matching Technique (PSMT)

In order to derive the partial effect of the highway on the highway development related outcome variables, it has already been mentioned that the Propensity Score Matching Technique has been used to identify the groups of matching households in the control zone for any given household in the influence zone for the sake of comparison of values of outcome variables between the influence zone and the control zone to estimate the impact of proximity on those variables. Originally the PSMT was suggested as a device for selecting the matching control units corresponding to an individual treated unit in a nonrandomised experiment set up for the estimation of the treatment effects of controlled experiments (Rosenbaum and Rubin, 1983). This technique has been found to be convenient for measuring the impact of socio-economic welfare programmes (Baker, 2000; Jalan and Ravallion, 2003). The literature, however, identifies the treatment group and the control group for such programme or project evaluation.

This technique involves two major steps – viz., the selection of matched unit(s) (that is, households here) corresponding to every individual participating unit of a sample of the participating units and then measuring the impact from the differences of the impact variables between the participating (treatment group) and the non participating units (control group). In our context, the influence zone and the control zone are analogous to the treatment group and the control group of the literature. The households located in the influence zone are deemed to be the project participants and those in the control zone as non participants (Asian Institute of Transport Development, 2003) in the application of PSMT in the case of road development project. Unlike in the case of socio-economic welfare project in education or health where the techniques have been used (Baker 2000 , Jalan and Ravillion 2003), the individual households have no freedom of choice to be participants or not as the highway and the household locations are already given. That is why we devise the categories of influence zone and control zone to identify the households who are deemed to be treated by making the accessibility to highway available. Those who are in the control zone are supposed to be lying outside the zone of such beneficiary of proximity to the highway.

The propensity score matching technique requires the estimation of a logit or probit model. The logit model would yield the estimate of the probability that any randomly chosen household of the sample which is characterised by a set of factors or covariates would fall in the influence zone[10]. For any given household in the influence zone, the matching households in the control zone are defined to be the ones that are the closest to it in terms of such estimated probability (P) for the concerned household unit or in terms of its odd ratio $\frac{P}{1-P}$. The simplest matching method is to find the five nearest neighbours in terms of the probability P. Alternatively, the matching groups may consist of those household units for which the odd ratios lie in a defined interval around the value of $\frac{P}{1-P}$ for the concerned influence zone unit. In the exercise of this study the matching criterion was taken to be $|\pi_i - \pi_j| < 0.0707$ where $\pi_i = l_n \frac{P_i}{1-P_i}$ and i and j are the suffixes of the sample households being

[10]For the present exercise around 40 village and household level characteristics (like proportion of scheduled caste and scheduled tribe population, existence of market, health care, education facility, distance from nearest urban centre, availability of electricity etc. and household size, social class, occupational characteristics, other infrastructural connectivities, etc.) were used as explanatory variables of the fitted logit model. The estimated mean values of each of these variables for the households of influence zone and control zone compared to ascertain that PSMT worked satisfactorily.

matched between the two zones. The difference between the actual value of the outcome variable in the influence zone and the mean value of the same for the matched control group would yield the estimate of the gain due to the programme or the project for that observation[11].

A few pertinent issues about the appropriateness of impact analysis based on PSMT for a road-related project like the present one should be mentioned here. These range from the question of feasibility of finding a satisfactory control zone to the applicability of PSMT for finding the matching control zone households based on the estimated propensity score. Consider first, the issue of finding a satisfactory control zone. NH2 being essentially a *historical* road connecting the country's northern and eastern parts, the villages located along NH2 within its proximity are in some kind of a long run social equilibrium state and therefore these villages (that is, those falling in the influence zone) will be, in general, non-comparable with those along other routes and in the interior. Whether PSMT based on a few observable characteristics would succeed in identifying appropriate matched control sample households is a valuable empirical question. The question of appropriateness of PSMT in the present context also arises from another point. As mentioned above, PSMT would be appropriate in the case of a project with voluntary participation where the participation decision depends on a set of relevant household variables. In the case of a road-related project, however, the participation is non-voluntary as all the households of the influence zone have to be regarded as the participating households and those of the control zone as the non-participating households. Thus, the notion of participation here

[11]For further details of application of the PSMT in the context of the present study one may refer to the AITD (2003).

is thus artificial.

4.3. Method of Differencing

As already mentioned, the method of differencing has been used in the present exercise to estimate the *partial* effect of NH2 in respect of individual impact/outcome variables. Briefly, this method is as follows. Consider a specific impact/outcome variable and let $y_i$be the observed value for the ith sample household of the influence zone, $y_j^i; j = 1, 2, ..., n_i$ the observed values for the corresponding $n_i$matched sample households of the *control* zone and $\bar{y}^i$ be the average of y_j^i $j = 1, 2, ..., n_i$ using the household sizes of the matching households of the control zone as weights[12]. Thus, $(y_i - \bar{y}^i)$is a measure of the impact of NH2 in terms of the specific variable considered for the concerned household of the influence zone. In terms of the specific variable, an overall measure of the impact in absolute terms for any variable would be $(\bar{y}_{IZ} - \bar{y}_{CZ})$or $\left(\frac{(\bar{y}_{IZ} - \bar{y}_{CZ})}{\bar{y}_{CZ}}\right) \times 100$ in percentage form, where $\bar{y}_{IZ}$ and $\bar{y}_{CZ}$ are appropriate averages of the y_i's and $\bar{y}^i$'s over all the sample households of the influence zone using the size of the household of the influence zone as the weights.

The impact measurement procedure described above is known as the *single difference method.* This method will measure the partial impact of an existing highway (based on the baseline survey data, as done in the present exercise). To measure the impact of a new highway or widening of an existing highway, one would need to compare the levels of impact/outcome variables of the

[12]There are many alternative ways of working out the average value of the impact variable for the matched households ranging from the use of the nearest neighbour weights to the nonparametric weights based on kernel function. For details, see Heckman et. al. (1997) and also Jalan and Ravallion (2001, 2003).

pre- and post-project periods. The relevant method, known as the *double difference method*, that would measure the partial effect of such a project is briefly as follows. Let $D_1 = (\bar{y}^1_{IZ} - \bar{y}^1_{CZ})$and $D_2 = (\bar{y}^2_{IZ} - \bar{y}^2_{CZ})$be differences in the mean levels of an outcome variable of the participating households (that is, influence zone households) and their matched households in the pre- and post-project periods (estimated from the baseline and the post-project survey data), respectively. The estimate of the impact of the project (that is, widening of the highway) is $DD = D_2 - D_1$[13].

4.4. The Data

In view of the above discussions, it is now clear that the present study contains two substantive empirical parts. The first part deals with an empirical delineation of the *influence* zone of the highway. Given this, the second part is concerned with the estimation of the impact of the highway on the rural households living in the influence zone. These have been done on the basis of a set of cross-section household level data on a large number of relevant variables collected through a pre-project base line survey relating to the four laning of the highway NH2 which was conducted during September-November, 2002. A total of 3200 sample households from 200 sample poverty stricken villages located within a horizontal distance of up to 7 km on either side of the 955 km stretch of the NH2 between Agra in Uttar Pradesh and Dhanbad in Jharkhand were covered in this survey.

[13]Note that one may write alternatively $DD = (\bar{y}^2_{IZ} - \bar{y}^1_{IZ}) - (\bar{y}^2_{CZ} - \bar{y}^1_{CZ})$, which measures the difference in the mean levels of the outcome variable between the pre- and post-project situations of the influence zone households and their matched control zone households. Either way this estimation would need data collected from the same set of sample households in the pre- and post-project surveys – a requirement which may prove to be a demanding one.

The survey design was two-stage stratified probability sampling with the village and the household as the first and second stage units, respectively. Factors such as agro-climatic characteristics, incidence of poverty, distance from the highway were used for the stratification of the universe[14]. Out of the 1697 villages located in the neighbourhood of the two sides of the highway in the rural segments of the 900km distance between Agra and Dhanbad , the sample covered 200 villages, 3200 households of such villages whose total household membership was 20,389. This represented coverage of 11.78% of villages, 1.26% of households and 1.15% of population of the concerned area which represented substantially higher coverage than that of the household surveys of the National Sample Survey Organisation Government of India with reference to the population set. Of these 3200 households, the numbers of influence zone and control zone households have been 2112 and 1088 respectively.

The data set covers household level information on a number of variables. These relate to the household level demographic, social and economic characteristics, employment and economic activities, land and asset holding, farm and non-farm production, income, consumption of goods and services, access to and use of education and health services and a detailed travel diary. While we have considered the access to health facility as the beneficial impact of high way development, there can as well be effects of negative externalities of such development like incidence of HIV/AIDS and air pollution related health effects. We have not, however, considered in the present study these negative externalities due to some environmental and disease specific health

[14]See Asian Institute of Transport Development (2003), for a detailed description of the Survey Design.

data related issues for the rural stretches of the high way covered in the survey. The travel diary generated from the survey data gives the records of purpose, mode of transport used, time required and cost of all trips and transportation made by the household members during a seven day reference period. Based on this, 30 impact or outcome variables have been defined. A list of the chosen set of impact or outcome variables classified into seven groups is given in Table 1.

5. The Results

The outcome variables as listed can be classified into seven groups of household level outcomes for the purpose of the present analysis. These relate to (i) the poverty status, (ii) mobility, (iii) income, employment and occupation, (iv) housing condition and asset ownership, (v) access to health, education and other infrastructural facilities, (vi) attitudinal variables reflecting a household's perception about its own poverty status and the possibility of improvement of its employment opportunities due to the widening of NH2, and, finally, (vii) composite indices of household well-being that combine relevant outcome variables. The total number of outcome variables considered is 30.

5.1. Comparison of Means

As a preliminary analysis, the sample mean values of the outcome variables for the influence and control zones have been compared. Table 2 presents these mean values and also the percentage gain in well-being related variables in the influence zone (IZ) over control zone (CZ) by simple comparison of the means without using any PSMT. It may be noted that out of 30 variables, for 15 variables the means are significantly different. The set of variables for which the difference in means have turned out to be significant includes the variables relating to poverty status (H1, H2), three of the ten variables relating to mobility (H4, H7, H8), three of the seven variables relating to income and employment (H15, H16, H19), the variables relating to asset ownership (H20, H21, H22) and attitudinal response (H26, H27) and, finally, two of the three well-being indices (H29, H30). These results, thus, on the whole, support the basic hypothesis that proximity of NH2 gives a better socioeconomic well-being, on the whole.

5.2. The Results of Non-Parametric Regression Analysis (NRA)

A working hypothesis running through the present exercise has been the gradient of change postulate with reference to the distance and the NRA results provide a convenient and objective procedure for empirically verifying this postulate. The NRA results would also supplement the results of impact estimation based on PSMT presented later[15]. A summary of the NRA results is presented in Table 3. Appendix Figures 1 to 8 present the graphs of eight selected variables illustrative of such NRA results.

As Table 3 would suggest, on the whole the NRA results support the gradient of change hypothesis, showing a systematic relationship with distance from NH2. However, it may be noted that the distance from NH2 may not be the only explanatory variable influencing the set of outcome variables.

[15]There is, however, a basic qualitative difference between the PSMT and the NRA procedure. The impact measured for the individual outcome variables by the PSMT is, in principle, an estimate of the pure partial effect of the proximity to the NH2 (because of the use of matched sets of households while computing the impact). The NRA, on the other hand, measures the total rather than the partial effect of proximity to the NH2 on the variable concerned.

There are other variables like the household and the village-level characteristics that together with the distance variable affect these outcome variables. Most of the regression curves show a change in gradient indicating change in the structural pattern of relationship in the range of approach distance of 4 to 5 km from NH2.

5.3. The Results of the Propensity Score Matching Techniques (PSMT).

As already explained, PSMT involves two steps. In the first step, a binary logit analysis is performed (based on the entire sample) to estimate the propensity scores for individual sample households. This logit model considers variables characterizing the socio-economic conditions defined both at the household and the village level. The results of this logit regression model used for identifying the matching sample households in the two zones are presented in Table 4. Using this estimated model, a set of matched control zone households for each influence zone household has been identified and for every outcome variable, the averages of observed values for the influence zone sample and the corresponding control zone *matched* sample were computed using the household sizes as the weighting factor. The difference between these two computed averages for an outcome variable is taken as the estimated impact of NH2 in respect of the particular variable. The estimated percentage gain (loss) for individual outcome variables based on single differencing is presented in Table 5. The negative values indicate loss in the attainment of the outcome variable values in the influence zone (IZ) when compared with those of the control zone (CZ). For a normal outcome variable for which a higher numerical value indicates betterment of socio-economic well-being for the influence zone households (for example, per capita income) when compared with the matched households of the control zone, the gain due to proximity to NH2 has been measured by as$(\frac{(\bar{y}_{IZ}-\bar{y}_{CZ})}{\bar{y}_{CZ}}) \times 100$.

However, there are poverty related outcome variables for which the values are expected to be numerically smaller to indicate betterment of socio-economic condition in the influence zone. For such an outcome variable the gain due to proximity of NH2 has been measured as$(\frac{(\bar{y}_{CZ}-\bar{y}_{IZ})}{\bar{y}_{CZ}}) \times 100$,where the symbols are as already defined. Any negative value or loss if reported in Table 5 thus represents the change to be not in the expected direction. A comparison of results of the gain in the values of outcome variables before the use of propensity score matching technique (PSMT) as given in Table 2 and those obtained as estimates of such gains by using PSMT as given in Table 5 would indicate how effective the matching of households has been in separating the effects of proximity to highway from other factors.

As Table 5 shows, the estimated gain due to proximity of NH2 is positive except for a few variables. It may also be noted that it is only for the poverty status related three variables – H1, H2 and H26 for which a lower numerical value would represent gain in socio-economic well-being. The variables for which an unexpected negative impact is estimated are the poverty status measured in terms of poverty line based on the monthly per capita consumption expenditure MPCE (H2), per capita trip rates for education and health (H6, H7); per capita consumption expenditure (H14); and access to educational facilities[16] (H23, H24).

[16]However, these results may not appear to be unrealistic, if one recognises the fact that the scope for child employment is likely to be greater in the neighbourhood of the NH2.

5.4. Developmental Implications of Proximity to NH2

The results of NRA and PSMT-based single differencing analysis (PSMT-SDA) based on household level data may be put together to see the possible developmental implications of proximity to NH2 of a household. In what follows, we summarise how the specific aspects of mobility and well-being considered in the present exercise are seen to have been affected by the proximity to NH2. While one compares the NRA results or difference between the mean values of a variable between the influence zone and the control zone, with the PSMT-SDA results one has to remember that the latter takes difference between the averages of values for the matched units, while the former ones do not involve any such grouping for comparison based on matching and represent the differences between the overall ratios or the means of the entire sub-samples of the two zones. The discrepancy in the qualitative nature of the results as per the two approaches will have to be interpreted accordingly without necessarily implying inconsistency of results.

5.4.1 *Poverty Status*

For obvious reasons, the impact of a public investment project on poverty is considered to be of utmost importance in a poverty stricken country like India. As already argued, one should expect a gain in poverty reduction for greater proximity to the NH2 because access to NH2 would promote income generation by stimulating economic activities through various linkage effects. Two different outcome variables relating to the poverty status of a household considered here are defined in terms of a poverty line based on the monthly per capita income MPCY (H1) and the monthly per capita consumption expenditure MPCE (H2) as estimated from the sample data. In view of the currently on going debate on poverty status in India, we would like to mention here that we have taken the state level poverty lines for Uttar Pradesh, Bihar, Jharkhand for ascertaining the poverty status of the individual households[17].

As the PSMT-SDA results show, proximity to the NH2 results in a gain in terms of poverty reduction based on the MPCY, but not for the other poverty variable. In the case of the MPCE-based poverty measure, a very small loss is estimated. The NRA result shows the expected inverse relationship with distance from NH2 for the poverty measure based on the MPCY over the entire 0-8 km distance range. However, for the other poverty variable estimate, an inverse relationship is found to be valid only up to a distance of 3.5 km as per the NRA result.[18]

[17]Whether consumer expenditure or income should be used for poverty measurement has been a matter of intense debate. The use of consumer expenditure data for this purpose is preferred for two reasons: (i) consumer expenditure relates more directly to nutritional deprivation, a major dimension of absolute poverty in developing countries, and (ii) compared to income, consumer expenditure data may be more reliable as a measure of well-being due to greater recall lapse and the tendency to underreport income. For the MPCE-based poverty variable (H2), the official poverty line has been adopted. The poverty lines for Uttar Pradesh, Bihar and Jharkhand as given by the Planning Commission with appropriate correction for price changes have been used. These poverty lines are Rs.357.73 for Uttar Pradesh, Rs.333.07 for Bihar, and Rs.333.07 for Jharkhand at 2002-2003 prices. The poverty line in terms of the MPCY has been estimated by inverse projection. Using the observed data on MPCE and MPCY, a log-linear engel curve has been estimated. The MPCY corresponding to the given poverty line in terms of MPCE of the survey has been calculated from the fitted engel curve. The poverty lines in terms of income thus estimated are Rs.377.77 for Uttar Pradesh, Rs.332.53 for Bihar and Rs.332.53 for Jharkhand at 2002-2003 prices.

[18]Beyond 3.5 km, the relationship of H2 with the distance from the NH2 is found to be wavy.

5.4.2 *Mobility*

As mentioned earlier, the direct economic benefits of proximity to the NH2 are expected to arise mostly out of enhanced mobility of the population living near this highway. A household living closer to NH2 should have greater movement for the purposes of travel for work, business, education and health and thus have a larger per capita trip rate. The analysis of the impact on mobility has been, therefore, an important element of the present study.

Of the ten mobility-related outcome variables that have been considered (viz., variables H3-H12), the PSMT-SDA results indicate gain for all except per capita trip rates for education (H6) and health services (H7). The corresponding NRA results are qualitatively similar, but appear to be somewhat more insightful. For example, the nonparametric regression graphs show that for all the variables relating to mobility the relationship with the distance from the NH2 is a monotonic inverse one, either throughout the entire range of distance or up to a certain distance. The NRA results show an inverse relationship with distance of per capita trip rate involving travel on NH2 (H8) and a positive relationship is observed for the variable travel time on NH2 (H11). Finally, for most of the variables the NRA graphs display a change in curvature around a distance of 4-6 km.

5.4.3 *Income, Employment and Occupation*

The PSMT-SDA results show a gain due to proximity to NH2 for all the variables in this group except for the per capita consumer expenditure (H14). The NRA results also support this. The NRA results further show that while per capita household income tends to decline with distance from the NH2, the consumption expenditure is a monotonically declining up to a distance of about 3.5 km, beyond which the relationship is not so well defined. In respect of the share of household income from self-employment in the non-farm sector (which is an important indicator of rural development), the estimated nonparametric regression curve shows an inverse relationship with distance from NH2 up to a distance of 5.5 km. This result is quite expected as more non-farm activities generally develop in the vicinity of a highway.

The NRA results show further that the labour participation rate for the population in the 15-59 age group, overall as well as that for the female population, tends to decline between 2 and 8 km from the NH2 after an initial rise in the 0-2 km distance range. The PSMT-SDA results are consistent with this observed pattern of relationship. So far as the impact on the occupational pattern is concerned, the NRA results show that the share of employment in non-agricultural activities declines up to a distance of 7 km from the NH2. The corresponding PSMT-SDA result also shows a gain for this variable in the influence zone over the control zone.

5.4.4 *Assets and Ownership*

As the PSMT-SDA results show, the proximity to the NH2 leads to a gain for all the outcome variables of this group. The NRA results indicate that the probability of a household being landless tends to decline with distance from NH2 between 2 km and 8 km after an initial rise in the range of 0-2 km and broadly corroborate the corresponding PSMT-SDA results. The observed inverse relationship for the proportion of landless with distance from the highway may be explained in terms of the corresponding inverse relationship between the land price and the distance from the NH2. As regards the ownership of a (information or media related) consumer durables (for example, radio, TV, etc.) an inverse relationship with the distance with a gentle wavy pattern is obtained from the NRA results. The estimated relationship for the variable relating to ownership of at least one motorised transport is observed to be clearly an inverse one up to a distance of 6.5 km. The asset ownership pattern thus seems to overall conform to the hypothesis that the households located nearer the NH2 are better endowed with the kind of assets considered here. However, since the asset ownership has been considered here in a very limited sense, the findings may be treated only as indicative.

5.4.5 *Access to Education and Health Facilities*

Three outcome variables relating to the access to education and health facilities have been considered in the present exercise. Out of these, it is only for the variable measuring proportion of household members visiting medical personnel during last 6 months (H25), the PSMT-SDA result shows a gain due to proximity to NH2. The variables for which a loss due to proximity to NH2 has been observed are the proportion of school-going children in the household (H23) and that of female school-going children (H24). The nonparametric regression curve for the proportion of school-going children is found to be positively related to the distance and that for the proportion of female school-going children is seen to rise up to 2.5 km and then decline. These results may not be unrealistic, if one takes into account the fact that scope for child employment is much greater in the neighbourhood of NH2. As regards the access to health facilities, the NRA results show that the relationship between the distance from the NH2 and the proportion of family members visiting medical personnel is an inverse one up to a distance of 5.5 km. This result may be suggestive of the possibility that the availability or access to health facilities declines as the distance from the NH2 increases. These findings are also in overall conformity with those relating to the mobility for education, though not entirely in respect of mobility for the health services.

5.4.6 *Well-being Indices*

Finally, three indices have been computed as indices of attainment of well-being and their behaviour estimated with variation in respect of proximity to the highway. The first one is based on the outcome variables capturing the attainment of overall well-being in respect of income, employment, and health. The second one is based on transport mobility and the third index captures the extent of access to basic amenities like electricity, safe drinking water, proper sanitation etc. The values of these indices have been compiled for the individual sample households using the available data and the effect of the proximity to the NH2 on these indices have been examined.

The overall index of well-being has been obtained by combining the indices for the variables of (i) per capita income, (ii) share of income earned from self-employment in non-agricultural activities, (iii) labour participation rate, (iv) proportion of family members visiting health personnel and (v) proportion of school-going children. It may be noted that some of these variables are related to a household's entitlement to well-being (like (i) and (iii)), while the others relate to the capability of the households for having access to well-being (viz., (ii), (iv) and (v)). The variable-specific indices have been constructed as BORDA index by first merit-ranking all the households according to their scores in respect of the value of the concerned variable and then calculating an ordinal aggregator as per the BORDA rule. The households have been ranked according to the sum of the rank scores on the constituent counts as per this rule and have been given the overall rank in a descending order.

The nonparametric regression curve for the overall well-being index has shown an inverse relationship with distance from the NH2. Consistent with this, the corresponding PSMT-SDA results have also indicated a substantive gain of 30.94 percent for the influence zone over the control zone because of the proximity to NH2.

The well-being index relating to transport mobility has been obtained using the data on the mobility-related outcome variables, viz., per capita trip rate for trips relating to work, visit to market, health facilities and education purposes by similar rule of aggregation as in the case of computing the overall well-being index. The NRA results show an inverse relationship of this index with distance from the NH2 and the PSMT-SDA result also indicates a gain for the influence zone over the control zone in this respect.

Finally, the well-being index relating to the access to amenities has been estimated by combining the household-specific rank scores based on the qualitative information relating to the type of dwelling, use of electricity, toilet facilities, access to drinking water and use of biomass-based energy[19]. As per the NRA results, this well-being index, like the other two, is found to be inversely related with the distance from the NH2 and the PSMT-SDA results also show a gain for the households in the influence zone over the matching group of households of the control zone.

[19]It may be noted that a lower merit rank score has been given as a household's dependence on biomass increases.

6. Conclusion

The basic premise that has been empirically investigated in this paper relates to the effect that a highway or its widening may have on the socioeconomic lives of the people, especially the poorer people, living in the proximity of the highway. From an aggregative developmental perspective, an improvement of transport infrastructure is expected to lead to income growth by promoting traffic and freight movement and thereby expanding the size and the access to the markets through a variety of direct and indirect linkage effects. The consideration of equity demands that the benefits of such huge public investment schemes trickle down to the poorest in the society. This would indeed be so if, at the household level, the partial effect of a highway or its widening ultimately leads to an improvement in the level of well-being of the households, particularly the poorer ones, living in the proximity of the highway.

As we have argued here analytically in terms of a *rural household model*, such positive effects at the micro level is likely to take place, because the closer a household is to the highway, the greater would be its mobility and connectivity and hence access to the various economic opportunities and amenities of life. To put it differently, one would expect the possible welfare effects of a highway on the households living in proximity to decline as the approach distance of the household from the highway rises and ultimately to disappear beyond a threshold distance. In other words, for two otherwise comparable households, there would be a differential welfare impact of the highway or its widening – one living in the *influence zone* of the highway and the other living in the control zone.

Essentially, this premise has been empirically verified here using a household level base line survey data pertaining to the project of widening of a stretch the NH2, one of India's oldest trunk routes. The measuring of the socioeconomic impact of a road-related public investment project, be it a new road or widening of an existing one, involves conceptual as well as operational problems ranging from the delineation of the beneficiary population to segregate the pure partial effect of the project from the effects of the other interventions on the set of relevant outcome variables. Such innovative use of methodology of impact evaluation based on the NRA, PSMT and SDA as attempted in this paper may be used for the impact analysis of similar road-related public investment projects.

Coming to the results, the partial effect of the NH2 as separated from the effects of other factors has been observed to be positive and significant for such major aspects of a household as mobility, poverty status, earning and employment opportunities, asset holding, access to education and health services and on the summary indices of transport mobility, access to infrastructural facilities and overall well-being. However, our attempt to find out the separated effect of proximity as distinct from other influencing factors, despite the comparison of values between the matched households in the two zones, suffers from the limitation that we applied only a single difference method based on the base year data and not a double difference method as the pre-existence of NH2 data do not exist. As a result the impact of proximity as reflected in the single differences of the observed values of the outcome variables is possibly not entirely free from the influence of other factors having impact on them. Besides, the domain of outcome variables as defined, does not include the pollution and HIV / AIDS related adverse impact of proximity or up-gradation of highway, and thus leaves out certain aspects of environmental sustainability from consideration.

An important aspect of this exercise relates to the empirical delineation of the influence zone of the highway. As the NRA results have brought out quite convincingly, the beneficial influence of the NH2 extends up to an approach distance of 5 km on either side of the road and the influence declines thereafter, thus supporting the gradient of change hypothesis of transport economics. Interestingly, the delineated influence zone gains plausibility as the approach distance of 5 km perfectly matches with the distance that one can traverse on foot in an hour or on a bicycle in half an hour in typical Indian rural condition, so that one living beyond generally tends to have lesser access to the highway.

To conclude, contrary to the traditional view that a national highway facilitates mainly the intercity travel and transport of goods, the paper clearly brings out that it is also an integral part of the road network serving the local rural economies through which it passes. The results presented in this paper thus lend support to the view that in the developing economies like that of rural India, a large public investment project on road infrastructure development, apart from its broader general equilibrium effects on the national economy, may help in ameliorating rural poverty and improving the socioeconomic well-being of the people living in its proximity.

Table 1: List of variables used to study impact of proximity to NH2 based on the household-level data
Incidence of poverty
H1* Whether household is poor based on poverty line measured in terms of monthly per capita income (MPCY)
H2* Whether household is poor based on poverty line measured in terms of monthly per capita consumption expenditure (MPCE)
Mobility (weekly)
H3 Per capita trip rate
H4 Per capita trip rate for work
H5 Per capita trip rate for marketing
H6 Per capita trip rate for education
H7 Per capita trip rate for accessing health-related services
H8 Per capita trip rate for trips involving travel on NH2
H9 Per capita trip length for trips involving travel on NH2
H10 Per capita travel expense for trips involving travel on NH2
H11 Per capita travel time for trips involving travel on NH2
H12 Travel cost per person km for trips involving travel on NH2

Income, consumption, employment and occupation
H13 Per capita income (Rs./365 days) **H14** Per capita total consumption expenditure (Rs./30 days) **H15** Share of income from self-employment in non-agricultural activities **H16** Share of food in total consumption expenditure **H17** Proportion of working members in a household in age group 15-59 years **H18** Proportion of working female members in a household in age group 15-59 years **H19** Proportion of non-agricultural workers in total working household members.
Asset ownership **H20*** Whether a household is landless **H21*** Whether a household owns at least one information related consumer durable (like TV, radio, etc.) **H22*** Whether a household owns at least one motorised transport vehicle
Education and health **H23** Proportion of school-going children among all children in the household in age group 6-14 years **H24** Proportion of female school-going children among all female children in the household in age group 6-14 years **H25** Proportion of household members who availed of medical facilities during last six months

Attitudinal response
H26* Whether a household rates itself as poor
H27* Whether a household expects improvement in employment situation after widening of NH2
Well-being index
H28 Index of overall well-being based on income, employment, health and education
H29 Index of transport mobility
H30 Index of access to infrastructural facilities, assets and amenities
* *These are qualitative binary variables.*

Table 2: Mean values of the selected socio-economic variables at household level separately for influence and control zone

Variable@		Mean value		%age gain due to proximity to NH2
		Influence zone	Control zone	
Poverty status				
H1*	Proportion of poor households based on poverty line measured in terms of MPCY	34.61	37.32	7.26**
H2*	Proportion of poor households based on poverty line measured in terms of MPCE	43.89	44.3	0.93**
Mobility				
H3	Per capita trip rate	0.89	0.81	9.88
H4*	Per capita trip rate for work	0.55	0.42	30.95
H5	Per capita trip rate for marketing	0.09	0.09	
H6	Per capita trip rate for education	0.21	0.26	-19.23

H7*	Per capita trip rate for accessing health-related services	0.01	0.02	-50.00
H8*	Per capita trip rate involving travel on NH2	0.42	0.24	75.00
H9	Per capita trip length for trips involving travel on NH2	3.75	3.33	12.61
H10	Per capita travel expenses for trips involving travel on NH2	1.17	0.93	25.81
H11	Per capita travel time for trips involving travel on NH2	18.24	15.82	15.30
H12	Travel cost per person km for trips involving travel on NH2	0	0	
Income, employment and occupation				
H13	Per capita income (Rs./365 days)	8223	7906	4.01
H14	Per capita total consumption expenditure (Rs./30 days)	556	598	-7.02
H15*	Share of income from self-employment in non-agricultural activities	14.84	8.31	78.58
H16*	Share of food in total consumption expenditure	40.37	37.59	7.40
H17	Proportion of working members in a household in age group 15-59 years	49.38	48.45	1.92

H18	Proportion of working female members in a household in age group 15-59 years	8.76	8.25	6.18
H19*	Proportion of non-agricultural workers in total working household members.	56.09	48.4	15.89
Asset ownership				
H20*	Proportion of landless households	46.59	42.1	10.67
H21*	Proportion of households owning at least one information related consumer durable	25.8	25.28	2.06
H22*	Proportion of households owning at least one motorised transport vehicle	7.24	7.54	-3.98
Education and health				
H23	Proportion of school-going children among all children in age group 6-14 years	82.73	84.76	-2.39
H24	Proportion of female school-going children among all female children in age group 6-14 years	78.18	79.2	-1.29
H25	Proportion of household members who availed of medical facilities during last six months	13.4	12.85	4.28

Attitudinal response				
H26*	Proportion of households who rate themselves poor or very poor	45.28	50.64	10.58**
H27*	Proportion of households who expect improvement in employment situation after 4-laning of NH2	75.52	76.19	-0.88
Well-being index				
H28	Index of overall well-being based on income, employment, health and education	1056.48	1017.74	3.81
H29*	Index of transport mobility	1625.27	1529.37	6.27
H30*	Index of access to infrastructural facilities, assets and amenities.	1604.21	1483.63	8.13
@ For variables marked by * the difference of the mean values of the two zones are tested to be significant at 5 percent level, ** appropriate signs being used to show the percentage gain				

Table 3: A Summary of the Non-Parametric Regression Analysis Results (NRA) Results

Variable		Observed gradient of change
Poverty status		
H1	Proportion of poor households based on poverty line defined in terms of MPCY	Rises monotonically throughout
H2	Proportion of poor households based on poverty line defined in terms of MPCE	Rises up to 3.5 km
Mobility		
H3	Per capita trip rate	Declines up to 4 km
H4	Per capita trip rate for work	Declines throughout
H5	Per capita trip rate for market	Declines between 2 and 5km
H6	Per capita trip rate for education	Rises up to 6.5 km
H7	Per capita trip rate for accessing health-related services	Minor declining trend up to 3.5 km then reverses sharply thereafter
H8	Per capita trip rate for trips involving travel on NH2	Declines throughout
H9	Per capita trip length for trips involving travel on NH2	Rises up to 7 km

H10	Per capita travel expenses for trips involving travel on NH2	Declines up to 4 km
H11	Per capita travel time for trips involving travel on NH2	Rises throughout
H12	Travel cost per person km for trips involving travel on NH2	Declines up to 4.5 km
Income, employment and occupation		
H13	Per capita income (Rs./365 days)	Declines throughout
H14	Per capita total consumption expenditure (Rs./30 days)	Declines up to 3.5 km
H15	Share of income from self-employment in non-agricultural activities	Declines up to 5.5 km
H16	Share of food in total consumption expenditure	Wide fluctuations
H17	Proportion of working members in a household in age-group 15-59 years	Declines between 2 and 8 km
H18	Proportion of working female members in a household in age-group 15-59 years	Declines between 2 and 8 km
H19	Proportion of non-agricultural workers in total working household members	Declines monotonically up to 7 km
Asset ownership		
H20	Proportion of landless households	Declines between 2 and 8 km

H21	Proportion of households owning at least one information related consumer durable	Declines up to 4.5 km
H22	Proportion of households owning at least one motorised vehicle	Declines up to 6.5 km
Education and health		
H23	Proportion of school-going children among all children in age-group 6-14 years	Rises up to 8 km
H24	Proportion of female school-going children among all female children in age-group 6-14 years	Rises up to 2.5 km followed by decline
H25	Proportion of household members who availed of medical facilities during last six months	Sharp decline up to 5.5 km
Attitudinal response		
H26	Proportion of households who rate themselves poor or very poor	Rising trend up to 8 km
H27	Proportion of households who expect improvement in employment situation after 4-laning of NH2	Rising trend up to 5.5 km
Well-being index		
H28	Index of overall well-being based on income, employment, health and education	Declines sharply between 2 and 8 km

H29	Index of transport mobility	Declines Sharply between 2 and 5 km
H30	Index of access to infrastructural facilities, assets and amenities.	Declines up to 4.5 km

Table 4: Estimated logit model for determining probability of a household falling in the influence zone		
Explanatory variable	**Coeff-icient**	**p value**
Village variables		
Proportion of scheduled caste and scheduled tribe population in total population	0.00455	0.03
Average land resource base including common land per household	-0.20577	0
Whether there is a public health services facility	0.70849	0
Whether there is a market	1.10277	0
Whether there is a village-level body of local self-governance	-0.38139	0
Whether there is a primary school	0.03035	0.77
Proportion of net sown area in total reported area	0.00241	0.32
Whether the village electrified	0.02205	0.85
Distance from nearest urban centre	-0.02743	0
Household variables		
Operational land holding (ha)	-0.00397	0.17
Whether the household covered in any government anti-poverty programme.	0.16755	0.19
Whether the household found government officials not-helpful	0.29759	0.01

Whether the household failed to send a child to school for shortage of fund	-0.25268	0.11
Whether biomass is used as main fuel for cooking	0.00459	0.96
Whether the household has access to safe drinking water	-0.2929	0
Whether the household head is illiterate	0.18692	0.11
Whether the household head studied up to middle level	-0.09477	0.48
Whether the household head studied up to primary level	0.1601	0.32
Whether a female household member got elected as an office bearer of local self-government	-0.30284	0.13
Whether the household has separate space available for livestock	-0.23158	0.01
Whether the household head is satisfied with the functioning of the local self-government	0.39695	0
Whether the household treats female members on a par with male members for availing of educational facilities	0.17304	0.15
Whether the household has electricity	0.72581	0
Whether the household owns a clock/ wrist watch	-0.17513	0.06
Whether the household belongs to non-tribal backward community	-0.00223	0.98
Household size	0.0009	0.95
Whether the household belongs to tribal community	-0.42919	0.13

Type of dwelling – floor, walls and roof made of non-durable materials	-0.17269	0.1
Type of dwelling - at least one of floor, walls and roof made of durable material	-0.04458	0.7
Type of dwelling – Thatched roof	-0.52539	0.01
Whether the household head was born in the village	-0.12803	0.61
Whether a child of the household does not attend school because he/ she is needed for work at home/field	-0.19841	0.25
Number of rooms in the dwelling	0.03517	0.15
Whether the household owns a bicycle	-0.12679	0.18
Whether the household avails of the Public Distribution System	-0.67521	0
Proportion of female among adult household members	-0.00377	0.23
Whether the household belongs to Hindu community	-0.7362	0
Whether the household head rates the road development activity as important	-0.23708	0.01
Whether the household head is male	-0.1395	0.61
Whether the household has access to sanitary toilet	0.30798	0.06
Constant	2.53394	0

Table 5: Percentage gain (loss) of selected socio-economic impact variables due to proximity to NH2 of the households estimated by PSMT and Single Differencing Method

Variable		**percentage gain (loss) due to proximity to NH2**
Poverty status		
H1	Proportion of poor households based on poverty line measured in terms of MPCY	17.07
H2	Proportion of poor households based on poverty line measured in terms of MPCE	-0.17*
Mobility		
H3	Per capita trip rate	9.14
H4	Per capita trip rate for work	31.54
H5	Per capita trip rate for market	2.48
H6	Per capita trip rate for education	-12.52*
H7	Per capita trip rate for accessing health related services	-5.8
H8	Per capita trip rate for trips involving travel on NH2	79.1

H9	Per capita trip length for trips involving travel on NH2	13.81
H10	Per capita travel expenses for trips involving travel on NH2	15.79
H11	Per capita travel time for trips involving travel on NH2	16.61
H12	Travel cost per person km for trips involving travel on NH2	1.74
Income, employment and occupation		
H13	Per capita income (Rs./ 365 days)	3.22
H14	Per capita total consumption expenditure (Rs./30 days)	-7.27*
H15	Share of income from self-employment in non-agricultural activities	68.34
H16	Share of food in total consumption expenditure	11.83
H17	Proportion of working members in a household in age-group 15-59 years	1.76
H18	Proportion of working female members in a household in age-group 15-59 years	0.72
H19	Proportion of non-agricultural workers in total working household members	14.36
Asset ownership		
H20	Proportion of landless households	8.09
H21	Proportion of households owning at least one information-related consumer durable	23.02

H22	Proportion of households owning at least one motorised transport vehicle	58.22
Education and health		
H23	Proportion of school-going children among all children in age-group 6-14 years	-2.67*
H24	Proportion of female school-going children among all female children in age-group 6-14 years	-2.51*
H25	Proportion of household members who availed of medical facilities during last six months	4.05
Attitudinal response		
H26	Proportion of households who rate themselves poor or very poor	20.67
H27	Proportion of households who expect improvement in employment situation after 4-laning of NH2	3.43
Well-being index		
H28	Index of overall well-being based on income, employment, health and education	30.94
H29	Index of transport mobility	6.51
H30	Index of access to infrastructural facilities, assets and amenities	22.21
* Indicates unexpected results in terms of gain of IZ over CZ due to proximity to the NH2.		

Figures

Figure 1:

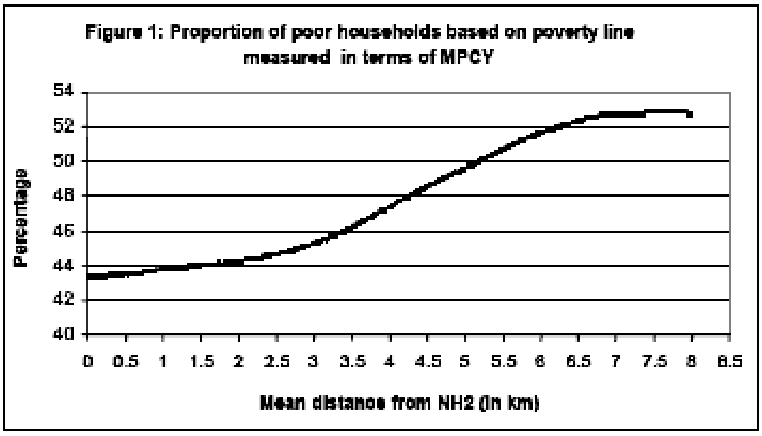


Figure 2:

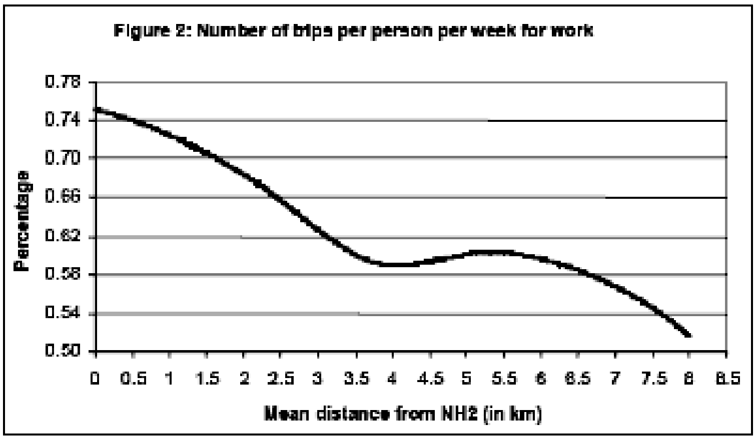


Figure 3:

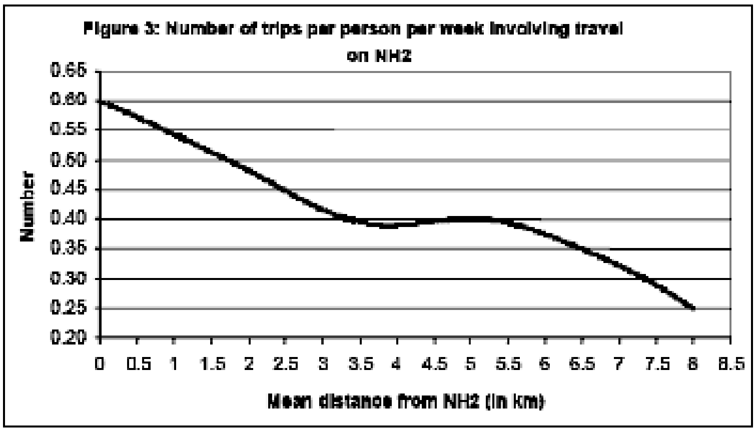


Figure 4:

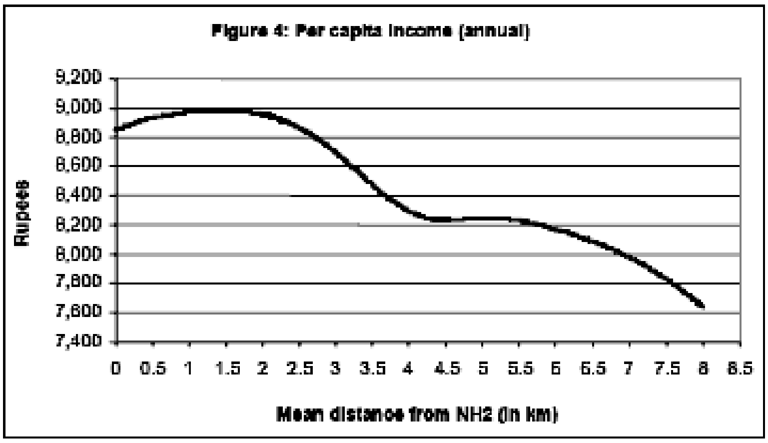


Figure 5:

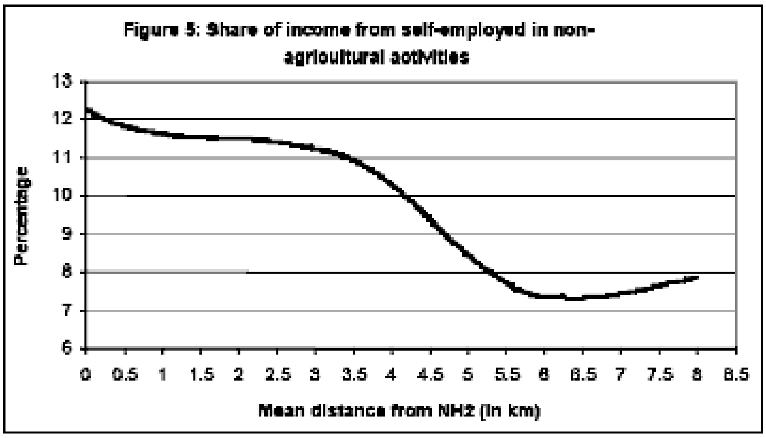


Figure 6:

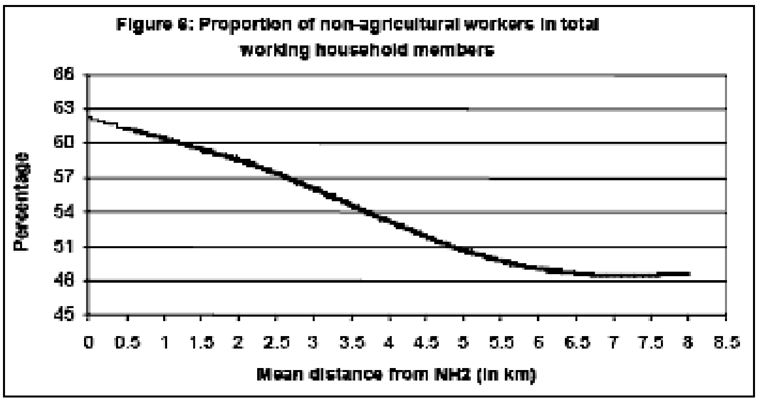


Figure 7:

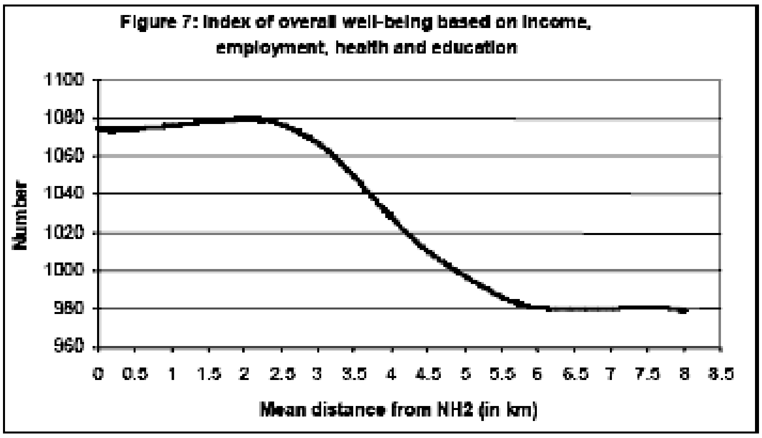
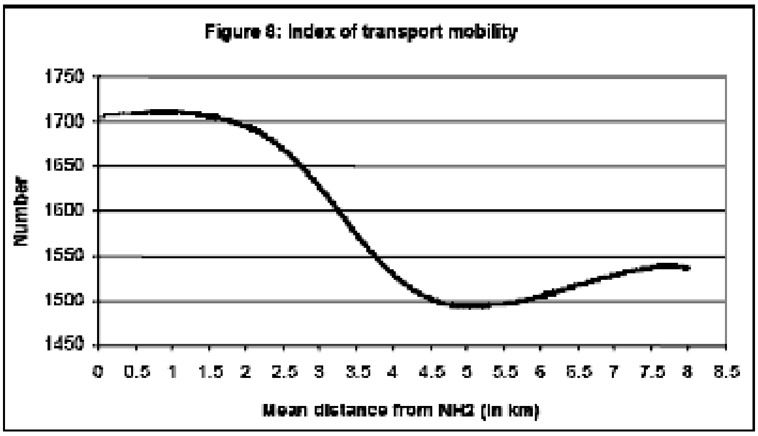


Figure 8:



Appendix 1

The Rural Household Model

Consider a representative rural household of the influence zone of the highway. We assume that the household makes rational choice in respect of allocation of the available time and the endowment of productive resources at its disposal so as to earn an income and spend it on current and future consumption to maximise household utility or welfare. Let the household's resource endowment consist of a given total available time (T) and given amounts of (agricultural) land (L) and capital stock (K) (like agricultural or other implements, irrigation facility, etc.) and assume that the household makes full utilization of K and L.

The household uses T for three purposes. Thus, $T_o$is used to work on *own account* (viz., for working on own farm or household enterprise), T_w is spent to work as a wage labourer (against a given market wage rate of w), T_t is spent on all kinds of travel together and the remaining time $T_{ls}(= T - T_o - T_w - T_t)$is enjoyed as leisure. Total travel time T_t, in turn, consists of time required for travel for (i) educational purposes(T_e), (ii) receiving health services(T_h), (iii) for moving goods and services for consumption purposes (T_c)and finally for (iv) productive purposes like commuting to work, etc. (T_L). The travel time for each purpose may be viewed as the product of *number of trips* made and the average time required per trip, so that the travel time functions may be specified as

$$T_j = n_j(C_j)\theta_j(\delta, \lambda) = T_j(C_j, \delta, \lambda); j = c, e, h, \quad (1a)$$

$$T_L = n_L(T_o + T_w)\theta_L(\delta, \lambda) = T_L(T_o + T_w, \delta, \lambda), \quad (1b)$$

where $n_j(.)$:number of trips per unit time and $\theta_j(.)$:time required per trip for the jth purpose, $j = c,\ e,\ h$ and L denoting travel connected with consumption of goods and services, educational services, health services and productive activities, respectively; C_j: quantity consumed of the jth type consumption, $j = c,\ e,\ h$; δ :approach distance of the highway from the household[20], λ :whether the highway has been widened[21]; and $T_o + T_w$:total time used for productive purposes. We assume that $T_j(.)$in 1(a) is increasing in $C_j$and $\delta$and decreasing in $\lambda$and $T_L(.)$in 1(b) is increasing in $(T_o + T_w)$ and $\delta$and decreasing inλ. The total time constraint faced by the household is thus

$$T = T_o + T_w + T_L(T_o + T_w, \delta, \lambda) + T_c(C_c,\ \delta,\ \lambda) + T_e(C_e,\ \delta, \lambda) + T_h(C_h,\ \delta, \lambda) + T_{ls} \quad (2)$$

The travel cost functions corresponding to (1a) – 1(b) may, in turn, be specified as

$$\tau_j = n_j(C_j)t_j(\delta, \lambda) = \tau_j(C_j, \delta, \lambda); j = c, e, h, \quad (3a)$$

$$\tau_L = n_L(T_o + T_w)t_L(\delta, \lambda) = \tau_L(T_o + T_w, \delta, \lambda), \quad (3b)$$

where t_j: cost per trip for the jth purpose[22]. In addition, there is a cost of transportation of goods for productive purposes, which is given by the transport cost function

$$\tau_p = \tau_p(Q, x, \delta, \lambda), \quad (3c)$$

where Q: (agricultural) output produced and x: vector of quantities of material inputs used for production including hired labour, if any.

The household income (Y) comprises income from three sources, viz.,

[20]It is assumed here that a trip for any purpose involves travel δ to reach the highway and then a travel of a fixed distance along the highway to reach the destination and only δ varies from household to household.

[21]This is essentially a dummy variable capturing the status of the highway.

[22]It may be mentioned that the travel time and the transport cost are often found to be strongly correlated with δ , the approach distance of the highway from the household. This may be because the market, educational institution, health centre are often located close on or to the highway.

mixed income from (agricultural) production activity (R) consisting of rent of land, profit on (own) capital employed and imputed wage of own labour, wage earning from off-farm employment (W) and remittances and other *exogenous* net transfer payments received (P). The *mixed* income R is given by

$$R = p_Q Q - p_x x - \tau_p(Q, x, \delta, \lambda), \tag{4}$$

where the (agricultural) output Q is given by the production function[23]

$$Q = Q(K, L, x, T_o, \mu), \tag{5}$$

μ being a parameter denoting the quality of land. In (4), p_Q and p_x are the given market prices of output and material inputs, respectively. The gross wage income[24] is given by $W = wT_w$, so that the aggregate income of the household is

$$Y = W + R + P \tag{6}$$

and the household's *full income* budget constraint is

$$Y = p_c C_c + p_e C_e + p_h C_h + \tau_c(C_c, \delta, \lambda) + \tau_e(C_e, \delta, \lambda) + \tau_h(C_h, \delta, \lambda) + \tau_L(T_o + T_w, \delta, \lambda) + S \tag{7}$$

where p_c, p_e, p_h are the given market prices of goods and services consumed, educational services and health services, respectively, and S denotes the savings made out of the current income.

Coming to the welfare or utility of the household, it is assumed that the household derives utility from the current consumption of goods and services, educational and health services as well as from the consumption in the next period made out of current savings. Taking these into account, a static household utility function is defined as

[23] We assume that the usual assumptions hold for this production function.

[24] This is being called the gross wage income, as it is before netting out the travel cost for commuting to the work place.

$u = u(C_c,\, C_e,\, C_h,\, T_{ls},\, \frac{(1+r)S}{p_f}\,),$ (8)

where $\frac{(1+r)S}{p_f}$ measures the value of consumption in the next period out of current period's savings, r and $p_f$being market interest rate and expected *composite* price of all purchases that may be made out of S in the next period. Here we assume $u(.)$ to be an increasing and concave function of the set of determining variables. We also assume that the discount rate of future consumption to be built into the utility function. Formally, thus the household's optimisation problem is the following.

Maximise the utility function (8) with respect to the set of variables

$T_o, T_w, T_{ls}, C_c, C_e, C_h, Q, S$ and x subject to the constraints defined by equations 1(a) – (7) [25].

The optimal values of the decision variables and the corresponding maximised value of the utility function *u(.)* will thus be determined in terms of the parameters of the system, viz.,$p_c,\, p_e,\, p_h,\, p_f,\, p_Q, p_x,\, r,\, w,\, \mu,\, K,\, L,\, P,\, \delta$ and λ. The set of impact or outcome variables (that is, the indicators of the level of socio-economic well-being or welfare of the household), defined over the domain of the optimal values of the decision/endogenous variables and some of the parameters of the model as explained in the main text, would ultimately be determined by δalong with other factors. Formally, let

$S_\delta\; = \{p_c, p_e, p_h,\, p_Q, p_x, w,\, r,\, p_f,\, K,\, L,\, P\}$ be the set of parameters (other than δ and λ)and

$M_\delta\; = \{\bar{C}_c, \bar{C}_e, \bar{C}_h, \bar{S}, \bar{Q}, \bar{x}, \bar{W}, \bar{R}, \bar{Y}, \bar{T}_j, j\; = o, w, L, c, h, e; \bar{\tau}_j, j\; = c, e, L, h, p\}$ be the corresponding set of optimal values of the endogenous variables. Let the set of impact or outcome variables defined as$Z_\delta \subseteq S_\delta \bigcup M_\delta$. The value of

[25]Under the usual assumptions of concavity of the preferences and of the production function, this constrained utility maximization problem will have a unique interior solution.

δ influences the values of the elements of $M_\delta$directly as well as S_δ indirectly through the general equilibrium effect of working of a more complete model for the economy. A shift in λ (that is, widening of the highway) would have similar effect on the impact variables through its direct effect on M_δ and indirect or secondary effects on S_δ as well as on M_δ, due to the induced variations in the price parameters.

References

1. Asian Institute of Transport Development (2003) Socio-economic Impact Evaluation of 4-laning of NH2 between Agra and Dhanbad on Rural Population, Report submitted to the NHAI, New Delhi.

2. Baker, J. (2000) Evaluating the Impact of Development Project on Poverty – A Handbook for Practitioners, Directions in Development Series, World Bank, Washington, D.C.

3. Deaton, A. and Dreze, J. (2002) Poverty and Inequality in India: A Reexamination, Economic and Political Weekly, September 7, 2002, pp. 3729-3748.

4. Heckman, J., Ichimura, H. and Todd, P. (1997) Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme, Review of Economic Studies, 64(4), pp. 605–54.

5. Härdle, W. (1990) Applied Nonparametric Regression, Cambridge University Press, Cambridge, U. K.

6. Jacoby, H. G. (2000) Access to Markets and the benefits of Rural Roads, Economic Journal, 110 (465), pp. 713-737.

7. Jalan, J., and Ravallion, M. (2001) Does Piped Water Reduce Diarrhea for Children in Rural India? World Bank Policy Research Working Paper No. 2664, World Bank, Washington, D. C.

8. Jalan, J and Ravallion, M. (2003) Estimating Benefit Incident for an Anti-poverty Program using Propensity Score Matching, Journal of Business and Economic Statistics, 21(1), pp. 19-30.

9. Levy, Herman (1996) Kingdom of Morocco – Impact Evaluation Report: Socio-economic Influence of Rural Roads, Operations Evaluation Department, World Bank, Washington, D.C.

10. Rosenbaum, P., and Rubin, D. (1983) The Central Role of Propensity Score in Observational Studies for Causal Effects, Biometrica, 70, 1, pp. 41-55.

11. Walters, A. A. (1968) A Development Model of Transport, American Economic Review, 58(2), pp. 360-378.