Elsevier Editorial System(tm) for Social Science & Medicine
Manuscript Draft

Manuscript Number:

Title: Estimating unbiased economies of scale of HIV prevention projects: a case study of Avahan

Article Type: Article

Keywords: economies of scale, efficiency, cost data, HIV prevention, scale-up, causal inference, Avahan

Corresponding Author: Dr. Aurelia Lepine,

Corresponding Author's Institution: London School of Hygiene and Tropical Medicine

First Author: Aurelia Lepine

Order of Authors: Aurelia Lepine; Anna Vassall; Sudha Chandrashekar; Elodie Blanc; Alexis Le Nestour

Manuscript Region of Origin: INDIA

Abstract: Governments and donors are investing considerable resources on HIV prevention in order to scale up these services rapidly. Given the current economic climate, providers of HIV prevention services increasingly need to demonstrate that these investments offer good 'value for money'. One of the primary routes to achieve efficiency is to take advantage of economies of scale (a reduction in the average cost of a health service as provision scales-up), yet empirical evidence on economies of scale is scarce. Methodologically, the estimation of economies of scale is hampered by several statistical issues preventing causal inference and thus making the estimation of economies of scale complex. In order to estimate unbiased economies of scale when scaling up HIV prevention services, we apply our analysis to one of the few HIV prevention programmes globally delivered at a large scale: the Indian Avahan initiative. We costed the project in the first four years of its scale-up. We develop a parsimonious empirical model and apply a system Generalized Method of Moments (GMM) and fixed-effects Instrumental Variable (IV) estimators to estimate unbiased economies of scale. We find that the scale-up of Avahan has generated high economies of scale suggesting that cost savings are possible when scaling-up HIV prevention in low and middle income countries.

*Cover Page

Title: Estimating unbiased economies of scale of HIV prevention projects: a case study of Avahan

Corresponding Author: Dr Aurelia Lepine, Aurelia.lepine@lshtm.ac.uk, +442079272845

Corresponding Author's Institution: London School of Hygiene and Tropical Medicine, 15-17 Tavistock place, WC1E 7HT, London, UK.

Order of Authors: Aurelia Lepine; Anna Vassall; Sudha Chandrashekar; Elodie Blanc, Alexis Le Nestour

Keywords: Avahan, efficiency, HIV prevention, scale-up, cost data, India.

JEL classification: C33, C36, D2, I1.

Funding: This work was funded by the World Bank.

Estimating unbiased economies of scale of HIV prevention projects: a case study of Avahan

Abstract (208 words)

Governments and donors are investing considerable resources on HIV prevention in order to scale up these services rapidly. Given the current economic climate, providers of HIV prevention services increasingly need to demonstrate that these investments offer good 'value for money'. One of the primary routes to achieve efficiency is to take advantage of economies of scale (a reduction in the average cost of a health service as provision scales-up), yet empirical evidence on economies of scale is scarce. Methodologically, the estimation of economies of scale is hampered by several statistical issues preventing causal inference and thus making the estimation of economies of scale complex. In order to estimate unbiased economies of scale when scaling up HIV prevention services, we apply our analysis to one of the few HIV prevention programmes globally delivered at a large scale: the Indian Avahan initiative. We costed the project in the first four years of its scale-up. We develop a parsimonious empirical model and apply a system Generalized Method of Moments (GMM) and fixed-effects Instrumental Variable (IV) estimators to estimate unbiased economies of scale. We find that the scale-up of Avahan has generated high economies of scale suggesting that cost savings are possible when scaling-up HIV prevention in low and middle income countries.

Keywords: economies of scale, efficiency, cost data, causal inference, HIV prevention, scale-up, Avahan.

JEL classification: C33, C36, D2, I1.

Research highlights:

- Estimation of economies of scale is hampered by several statistical biases that prevent from estimating unbiased economies of scale.
- When using appropriate empirical strategies to correct for these biases, we find that scaling-up HIV prevention interventions generate cost savings in low and middle income countries.

Main text (8312 words)

1. Introduction

The UNAIDS investment approach for an effective response to HIV/AIDS proposes the scale-up of HIV prevention for key populations as one of its core interventions. However, resources for expanding HIV prevention to all who may benefit for it remain scarce. Due to the recent flat-lining of development assistance for health; increased attention has been placed on identifying efficiency gains in HIV prevention in low and middle income countries. Recent global resource estimates of HIV prevention are based on the assumption that efficiency of services can improve with scale-up through economies of scale (Schwartländer et al., 2011). Yet, little is known about the existence and strength of these. This paper therefore aims to fill this gap by assessing the extent of the effect of scale on average cost of HIV prevention to key populations; and by doing so, to quantify the economies and diseconomies of scale.

Avahan is one of the largest HIV prevention project in the world and it was funded by the Bill & Melinda Gates Foundation (BMGF). NGOs are provided grants by Avahan through state lead partners (SLPs) to build a relationship with key populations (female sex workers (FSWs) and high risk men who have sex with men/ transgenders (HR-MSM/TG)) in order to provide HIV prevention services. The package of HIV prevention services provided includes outreach through peers, behaviour change communication, condom distribution, clinical services for sexually transmitted infections (STIs), community mobilisation, advocacy and enabling environment activities. Each peer educator provided services to about 25-50 people, sharing prevention information, distributing supplies (condoms and lubricants) and providing referral for STI management. STI clinics followed standard protocols for STI management. Community mobilisation, advocacy and enabling environment activities varied across the sites and included the formation of self-help groups, various drop-in centre events, skills training, legal literacy workshops, police and stakeholder sensitization, crisis response teams and access to social entitlements. HIV prevention across all four states was guided by a common minimum programme. These included a set of implementation standards for technical and managerial areas, project milestones, a common management framework, and a common set of indicators. Beyond this, there was flexibility to adapt services based on local context.

In the 4 study states (Andhra Pradesh, Karnataka, Tamil Nadu, Maharashtra), the Avahan initiative was implemented by 138 NGOs, supported by 6 state level partners (SLPs) and by pan-Avahan capacity building partners (contracted by the BMGF, which also had a national level office at Delhi). SLPs provided technical assistance to develop programme strategies, developed communication materials, enhanced the expertise of NGO staff, provided supportive supervision and supported the purchase and distribution of commodities. At the national level, Avahan developed over-arching programme strategies and organised annual partners meetings to coordinate with Indian authorities. The national level office also developed and maintained a computerised monitoring and information system; provided financial oversight; and monitored programme evaluation. International and national technical assistance was primarily focused on enhancing the expertise to deliver STI services, improving interpersonal communication, and providing support for advocacy and community mobilisation.

Avahan achieved an exceptionally rapid pace of scale-up of HIV prevention services; going from a coverage of 22,000 persons covered in December 2003 to 280,000 persons reached per year in December 2007 (Bill and Melinda Gates Foundation, 2008). In total in the data we collected, we observe that 725,040 high-risk persons (female sex workers and their clients and men who have sex with men) were reached between 2004 and 2007, 177 million condoms were directly distributed by Avahan NGOs and 529,381 STI visits were provided. Extensive research has been conducted to evaluate the impact and cost-effectiveness of the Avahan programme. Pickles et al. (2013) reported a decline in FSW HIV prevalence and between 142 and 2092 FSW HIV infections averted per district, with two-fold to nine-fold more among FSW clients. Correspondingly, Vassall et al. (2014) found a mean incremental cost per HIV infection averted of US$785 and a mean incremental cost per DALY averted of US$46. Future anti-retroviral treatment (ART) cost savings over the lifetime of the FSW cohort exposed to Avahan were estimated to be over US$ 77 million.

Despite the policy interest in this area, to date there are very few papers examining the determinants of average costs of HIV services in low and middle-income countries. Some recent studies (Marseille et al., 2012; Menzies, Berruti, & Blandford, 2012; Rosen, Long, & Sanne, 2008) present evidence regarding the relationship between HIV treatment and hospital size. On a sample of Zambian hospitals, Marseille et al. (2012) find that when the number of patient-years of ART increases by 1, the average cost decreases by 0.23 per cent. Menzies et al. (2012) from a sample of 54 clinical sites in five African countries find that when patient volume is doubled (from 5,000 to 10,000 patients), the average cost decreases by 28%. Other studies examine the relationship between scale-up and cost of HIV prevention (Dandona et al., 2005; Guinness, Kumaranayake, & Hanson, 2007; Guinness et al., 2005; Kumaranayake & Watts, 2005;

Marseille et al., 2007) but these papers did not manage to quantify the extent of economies of scale; due to the small sample sizes of the data sets used.

During the scale-up of Avahan, we collected an extensive data set on the cost of Avahan from the 64 districts of the 4 following Avahan states: Andra Pradesh, Karnataka, Maharastra and Tamil Nadu. In total, 138 Avahan NGOs were costed over 4 years from 2004 to 2007. This is the largest dataset on HIV prevention costs available globally. Additionally, since we are interested in estimating economies of scale of the whole programme, above level costs (i.e. programme administration costs, programme communication costs, state level partner costs, BMGF level costs and Pan-Avahan capacity building partner costs) are included; which is something that has rarely been done before. The method to allocate above the NGO level costs is based on programme records, expenditures reports and interviews with BMGF Avahan and SLP staff.

Although this data set provides a unique opportunity to explore economies of scale, there are substantial methodological challenges in order to quantify the level of economies of scale. In fact, while some previous papers have been informative regarding the main drivers of average cost of ART and HIV prevention services, they fail to accurately establish a causal effect between scale and average cost (endogeneity bias). There are several reasons for why these are likely to occur in relation to the estimation of economies of scale. Firstly, it is conceivable that endogeneity biases may occur due to the omission of pertinent variables in the analysis. For instance, the average cost of NGOs (hospitals) may be attributable to the competence and effort of the manager and peer-educators (health workers). Other important drivers of cost may be related to the characteristics of the NGO location. For example, Integrated Behavioural and Biological Assessment (IBBA) data show that the typology of sex-workers (street-based, bar-based, brothel-based etc.) varies widely between districts (Ramesh et al., 2008). In peri-urban districts most of sex-workers are brothel-based (National AIDS Research Institute, 2005-2010), while in big cities such as Chennai, Bangalore, Mysore and Hyderabad, most of sex-workers are street-based – which may mean it is more costly to reach them. Prices and their variation over time may also differ between districts and are likely to affect average cost. The longitudinal nature of our data allows the use of a panel estimator with NGO fixed effects which accounts for NGO time-invariant characteristics that are likely to be correlated both with the NGO size and its average cost (such as location and district characteristics). While a panel data set allows to account for unobserved omitted variables that are time-invariant, a common issue with such longitudinal data is the representativeness of the sample. In the data we have collected, every NGO that entered in the programme was automatically costed, allowing having an exhaustive and thus representative sample of the Avahan NGOs over the period considered. Attrition rate was extremely low; only 4 NGOs were lost over the 4 year period, therefore the data is not likely to suffer from a selection bias.

Secondly, a further potential source of endogeneity is simultaneous relationship between NGO size and average cost. In fact, the NGO size (or scale) is expected to influence average costs, which can result in the presence of economies and/or diseconomies of scale. However, one could argue that average cost may also affect NGO size since one may expect that NGOs that have a lower average cost will be able to expand coverage to key populations more easily. If we ignore this simultaneity bias, then the coefficient associated to NGO size will be artificially overestimated, which would result in an overestimation of economies of scale. Controlling for all drivers of average cost can encompass this issue, but given that in reality it is not possible to control for all the determinants of average cost, the use of appropriate empirical strategies is required to infer causality and obtain an unbiased estimation of economies of scale.

Finally, a last source of endogeneity could be due to random measurement error. In fact, the scale is measured by the number of persons reached by the Avahan NGO. The number of persons reached was collected by via the NGO's routine monitoring system and could contain some random measurement error. Random measurement error on the dependent variable lead to an attenuation (or dilution) bias i.e. to an underestimation of the coefficient. In an extreme case, imagine that the scale contains high random errors, then the coefficient associated to average cost will be zero.

In addition to the endogeneity biases, there could be another source of bias. Since we are interested in estimating economies of scale of the whole programme, the estimation of NGO average cost requires to allocate above-level costs (i.e. programme administration costs, programme communication costs, state level partner costs, Bill and Melinda Gates level costs and Pan-Avahan capacity building partner costs) to Avahan NGOs. Although, we used several tools to allocate above the NGO level costs (review of programme records, expenditures reports and interviews with BMGF Avahan and SLP staff), the allocation of the above-level costs can have an effect in the estimation of economies of scale. We present the allocation of above-level costs in the descriptive statistics section to ensure that this is not likely to be an issue in our model.

In order to estimate unbiased economies of scale, we first use a panel estimator with NGO fixed effects to account for NGO time-invariant characteristics that are likely to be correlated both with the NGO size and its average cost. Then, in order to test if the correction of these different sources of endogeneity results in an over- or underestimation of the economies of scale, we use a system Generalized Method of Moments (GMM) and an Instrumental Variable (IV) approach. We find that the two methods lead to similar conclusion.

Overall, the results suggest that the activity of NGOs does not generate any economies of scale, however to conduct their activities, NGOs rely on large fixed programme costs. This explains why when we consider the total programme average cost to estimate NGO costs; we find high economies of scale. Yet programme average costs are often not properly taken into account in global resource requirements estimate for HIV (Schwartländer et al., 2011). Instead, it is common practise to use estimates from studies conducted at the provider level and apply a constant mark-up. A further important finding is that at the NGO and programme levels, an L-shaped curve was found to be more appropriate than a U-shaped curve in both cases, consistent with an absence of diseconomies of scale.

The remainder of the paper is organized as follows. In the following section, we outline the elements of the Avahan design that are relevant to the estimation of economies of scale. The third section provides our descriptive statistics. In Section 4, we outline the empirical specifications and section 5 describes the econometric results. Finally, interpretation, policy recommendations and concluding remarks are contained in the last section.

2. Avahan design and funding mechanism: implications for the analysis of economies of scale

Broadly the extent of economies of scale observed in HIV prevention programmes may be explained by the design of the services offered, the funding mechanism that may discourage or encourage the realisation of economies of scale and the extent of above service (programmatic) costs. Economies of scale are commonly found to result from the: (1) existence of fixed costs, (2) learning by doing, (3) lower input prices due to high bargaining power of firms and (4) opportunities for specialisation. The industry analysed in this paper, is however slightly different from other industries, primarily due to the fact that NGOs do not necessarily operate as profit-maximisers and they are facing a monopsony funder. Specifically, NGOs buy their inputs (i.e. condoms, Sexually Transmitted Infections (STI) kits, etc.) from the same provider at a fixed price; and thus the size of an NGO is not likely to affect input prices. Secondly, NGOs are funded to provide a set mix of services such as outreach services, STI treatment, condom distribution and community mobilisation (Wheeler et al., 2012) and have little freedom to specialise where they perform the best. Thus, in principle NGOs are most likely to experience economies of scale due to the presence of fixed costs and the learning by doing that they gain over time.

Economic theory predicts that the cost function of NGOs should have a U-shape, with the presence of economies of scale at first and then with diseconomies of scale after a certain level of scale is reached. If NGO average cost is made of fixed costs, then larger NGOs should be more efficient than the smaller ones since they may be able to spread fixed costs over many persons. On the other hand, the difficulty to reach new potential beneficiaries after a certain point and the management complexity increasing with scale may explain why smaller NGOs may have a lower average cost than the larger ones. Despite the sound theoretical basis, empirically, a number of studies that have analysed average cost function of hospitals suggest that the cost function may be more consistent with an L-shaped curve (Lave & Lave, 1970). In the case of Avahan, an L-shaped relationship between scale and average cost may also be hypothesised for several reasons.

Firstly, during the first year of Avahan, NGOs set their budgets (and staffing levels) on the basis of the number of key population members in the district estimated through mapping and various forms of size estimation (Blanchard et al., 2008; Verma et al., 2010). In subsequent years, budgets were set in reference to both the number of high-risk population estimated as well as in reference to the number of persons reached the previous year. Due to the challenges of mapping key populations such as female sex workers, NGOs may learn by experience that the number of estimated key population members is often down-biased. This underestimation may be due to several factors. Migration during special events (Devadasi Festival, Hijra) may be excluded from the estimated population, dependent on the time of the survey. Similarly, seasonal work may also affect the number of persons reached, which may not be captured when the mapping takes place. Additionally, hidden, hard-to-reach populations may not be picked up by mapping method. Although NGOs have an incentive to reach mapped populations, the project does not provide to give any additional financial incentive to reach hidden or unmeasured populations.

Secondly, key population size is volatile; it may be affected by exogenous shocks such as climate shocks, since floods and droughts are common in the area of the study. Information that NGOs have regarding the level of persons reached in *t-1,* that informs their budgets, may therefore not correctly predict the number of person reached in *t*. NGOs are thus making decisions regarding the level of staff and consumables in a context of high uncertainty. This uncertainty may result in NGOs overestimating the numbers of consumables and staff needed to be ready on a stand-by basis limiting the chance to observe diseconomies of scale.

Thirdly, as Avahan scales-up it may become less costly to reach high-risk population through a reduction of HIV-related (or HIV service seeking) stigma. Avahan services become trusted by the community over time, NGOs sensitise community about HIV/AIDS through mass sensitization campaigns, celebration of World AIDS day, candle light memorial events for those who died of AIDS, local advocacy activities with police and political leaders.

Finally, given that the funding were allocated to NGOs based on their expected results NGOs did not have an incentive to reach a certain level of persons where their marginal cost exceeded their "marginal revenue". Thus, we hypothesise that the very large NGOs we observe are the ones that managed to reach a high number of persons at a low cost; and that the NGOs that would have had diseconomies of scale by reaching a similar number of persons would never be observed.

Therefore, we anticipate that the incentives generated by Avahan's programme and funding design lead to the absence of diseconomies of scale.

3. Descriptive statistics

Costs were obtained from NGOs, state level partners and the Bill and Melinda Gates foundation and are described in Chandrashekar et al. (2010). Costs include recurrent costs (personnel costs, project building and operating expenses, travel expenses, STI supplies, monitoring cost, information education & communication, training, condom supplies and indirect expenses) and capital costs (rent, equipment, furniture, vehicle, initial training, insurance and deposits, and start up cost).

The majority of costs (73%) for Avahan were incurred above the NGO level. An understanding of how this cost was allocated to NGOs is therefore pivotal to any investigation of economies of scale. Should the above service costs be determined by the levels of persons reached by each NGO then effectively they would be a variable cost; and lower economies of scale would be observed. However, if the allocation was made proportionally to each NGO, effectively programmatic cost would be a fixed cost which would increase the level of economies of scale observed. Methods for allocating programmatic costs to NGOs are complex as actual allocation is often difficult to observe. Our methods for allocating costs above the NGO level to NGOs were derived using a combination of programme records, expenditures reports and interviews with BMGF Avahan and SLP staff. The first step was to allocate national level programme costs to each SLP. This was done first by allocating specific grants to each SLP, and then for general programme management costs by using expenditure reports and mapping estimates of the key population covered by each SLP (the method reported by BMGF staff to be used for budget/ grant allocation to each SLP). Thereafter SLP costs (including BMGF costs) were first allocated to specific activity areas (for example programme management and expertise enhancement) within the SLP. This was done primarily on the basis of the description provided in detailed salaries reports and expenditure records; and where the allocation was not clear, interviews with SLP staff were conducted. Thereafter an allocation criterion for each activity cost was applied to allocate the cost to NGOs. The criteria used were derived after extensive interviews with staff on how they allocated their time and resources amongst NGOs. In the main, the allocation criterion used was either an equal division of cost, or an allocation based on estimated population size covered by the NGO. This latter measurement does not necessarily measure true output of each NGO, as they did not always covered the entire population in need. However, this was the best information SLPs had to hand when allocated resources such as communication materials. For some activities, costs could be directly allocated as the expenditure records including this description. This latter situation particularly applied in the case of support and supervision costs where detailed travel records were often described. Items such as STI drugs management could also be directly allocated based on order levels.

At the NGO level, costs were disaggregated by activity and input type. Field visits and time-sheets were conducted in order to estimate the share of labour costs allocated to different NGO sub-activities (outreach, community mobilisation, etc.). Unpaid volunteer time was estimated by the amount of time spent on the project and calculated based on peer educator salary. Other donated goods, such as commodities were valued using market prices. Capital costs were annualised using a discount rate of 3% and were assumed to have a life of between 5 and 10 years. Economic costs were computed valuing donated goods to their market price. All costs are presented in US$ 2008 in Table 1.

Insert Table 1

Table 2 presents the average cost (total and at the NGO level only) per person reached per year. The total average cost (that includes national level support cost) in the sample between 2004 and 2007 is US$231 while NGO only average cost is US$62. National level support costs contribute a large proportion of average cost (73%). We therefore conduct our analysis at both the NGO and total cost levels to provide a full picture of the existence of economies of scale. Typically, analyses of economies of scale focus on size of NGO examining only service delivery costs. However, if total costs are assessed and given that national level support costs are primarily fixed in terms of NGO size, then the economies of scale may be greater.

Insert Table 2

Table 3 presents the descriptive statistics included as explanatory variables in the estimation. The scale measure, as defined by the number of high-risk persons reached, was obtained from routine monitoring systems and was non-missing for 125 NGOs. On average, a NGO has reached 1,869 persons over the 4 year period; however this increases by year, reaching only 793 persons on average in 2004 against 2,405 in 2007. Since the distribution of the scale measure is right skewed, the variable was log transformed.

Insert Table 3

The relationship between average cost, measured by the number of high-risk population reached and scale is represented in Figure 1. While, NGO average cost represents only a small proportion of total average cost, it is interesting to note that its relationship to scale has the same shape as the total average costs, confirming the absence of diseconomies of scale.

Insert Figure 1

4. Empirical Estimation

From the above, we hypothesise that the presence of diseconomies of scale cannot be assumed to exist. We began our empirical estimation by testing this assumption by comparing a logarithmic form versus a quadratic functional form. We found that the logarithmic fit explains a larger share of the variance than the quadratic fit. One may want to note that a reasonable reason for such finding comes from the fact that although the squared term is statistically significant at 1%, the minimum of the average NGO cost including and excluding programme costs is 6,995 and 5,595 high-risk persons reached respectively. This corresponds to the last percentile in both NGO average cost distributions as we observe only 4 and 3 NGOs respectively with diseconomies of scale. For this reason, a logarithmic functional form was used.

Below, we present a parsimonious empirical model in order to understand the relationship between scale and average cost and to estimate unbiased economies of scale.

General estimation

We derive an equation to estimate the average cost curve empirically:

$$Log(AC_{it}) = \delta + \delta_1 \log Y_{it} + \delta_2 q_{it} + u_i + e_{it} \quad (4)$$

where AC_{it} is the average cost, note that we analyse both the NGO average cost that excludes programme cost AC_ngo_{it} and the NGO average cost that includes programme costs (total average cost) AC_tot_{it} ; Y_{it} is the size of NGO i in year t and is proxied by the number of high-risk population reached by the Avahan programme, q is a proxy of the quality of the services provided to reached population. Given that we do not have any direct measure of the quality of outreach, we proxy the quality by the intensity of the outreach. We consider the intensity of the services provided as measured by the number of condoms distributed on average per person reached and the number of STI visits in the NGO clinic conducted on average for each person reached. Finally, u_i is the NGO specific effect and e_{it} is an error term.

Cost function in the very short run

We first analyse the effect of scale on cost without any control variables (i.e. by removing q_{it}) in the very short run by adding temporal effects v_t:

$$Log(AC_{it}) = \delta_0 + \delta_1 \, logY_{it} + u_i + v_t + e_{it} \quad (4.1)$$

The inclusion of time dummies for each year, effectively quantifies the economies of scale once the decisions regarding the level of inputs required for that year have been made.

Cost function in the long run

To explore the relationship between cost and scale in a longer run, we remove time dummies as presented in equation (4.2). Interpreting this relationship as the 'true' long run average cost curve, we implicitly assume that NGOs are operating at the minimum of the short run average cost curve.

$$Log(AC_{it}) = \delta_0 + \delta_1 \, logY_{it} + u_i + e_{it} \quad (4.2)$$

However, there are good reasons to hypothesise that some NGOs are not observed at the minimum of their cost function; particularly NGOs that have not managed to reach full coverage of the estimated high-risk population in their catchment area, in the first few years of their start-up. Figure 2 supports this hypothesis, suggesting that average cost is lower when the number of persons reached (PR) is greater than the number of estimated population (EP) or when NGOs have reached a full coverage i.e. when PR/EP*100>100. Therefore, as a better proxy for the relationship between average cost and NGO size in the long run, we also decide to run the estimation only for the sub-sample of NGOs that have reached the estimated number of high-risk population in the district.

Insert Figure 2

Thus in equation (4.3) we conduct the same estimate as in equation (4.2) only for the sample of NGOs for whom the number of persons reached is greater than the number of estimated population, in order to have a better proxy of the long run average cost.

$$Log(AC_{it}) = \delta_0 + \delta_1 \, logY_{it} + u_i + e_{it} \quad \text{if PR>EP} \qquad (4.3)$$

Causal inference

To further explore the causal effect of scale-up on the average cost, we use a system GMM estimator as indicated in equation (4.4) and an IV approach presented in equation (4.5).

The use of these methods was mainly motivated by the suspicion of reverse causality and between scale and average cost. There are two reasons to suspect the presence of reverse causality. Firstly, NGOs that have a high average cost may be less able to reach new population than NGOs that have a low average cost. Secondly, NGOs that manage to substantially increase the number of persons reached in year *t* may be rewarded by the funding agency through an increase in budget in *t+1,* and depending on the use of this additional budget it could result in an over- (if the additional budget increases technical efficiency) or underestimation (if it decreases technical efficiency) of the effect of scale on the NGO average cost. In both cases, we would observe that the effect of scale on average cost is biased.

Regarding the GMM estimator in (4.4), first it is important to point out that the lagged dependent variable was not statistically significant, justifying the fact that we use the

GMM in a non-dynamic panel. The choice of the system GMM estimator is motivated by the fact that it has been found to be more efficient than the first-differenced GMM (Blundell & Bond, 1998). The two step robust variant of the GMM estimator used with finite-sample correction derived by Windmeijer (2005) is applied. To avoid bias due to a too large number of instruments, the GMM creates one instrument for every control variable and lag distance, rather than one for each time period, variable, and lag distance as suggested in Roodman (2006). In the model, we have only 4 instruments, we thus avoid the problem of too many instruments described in Roodman (2006). As shows in equation (4.4), we do not have any strictly exogenous variable nor predetermined variable. For the endogenous variable Y_{it}, the second and third lagged values are used as instruments. Note that the instruments are first-differenced for use in the first difference equations and used for instrumenting the levels equations.

$$Log(AC_{it}) = \delta_0 + \delta_1 \, logY_{it} + u_i + e_{it} \quad (4.4)$$

for i = 1,..., N and t = 2,..., T

with $E[u_i] = E[e_{it}] = E[u_i \, e_{it}] = 0$

The motivation for the use of the GMM is that, given the context of Avahan, there is no obvious reason that the level of scale in the past may directly affect the current average cost. One may argue that average cost might be partially determined by past scale, which requires to further question how Avahan funds have been allocated to NGOs. Interviews with Avahan NGOs and funders have highlighted that Avahan funding that NGOs received is mainly based on the NGOs previsions regarding their scale in year *t* rather than by past scale. Then it is likely that past scale may have an impact on current scale that may in its turn affect current cost but it is unlikely that past scale may directly affect current average cost.

Finally, we use an IV approach in order to test the robustness of the results obtained from the system GMM. To obtain a consistent estimator, we assume the existence of z_{it} that firstly, satisfies the assumption that Cov(z_{it}, e_{it})=0, so that the IV is uncorrelated with the error term. The second requirement involves the relationship between the IV and the endogenous variable that Cov(z_{it}, $logY_{it}$)$\neq$ 0. Thus, the instrument z needs to be strongly correlated with the number of persons reached but uncorrelated with other unobservable factors captured in the average cost error term.

$$Log(AC_{it}) = \delta_0 + \delta_1 \log Y_{it} + u_i + e_{it} \quad (4.5a)$$

$$Log(Y_{it}) = \alpha_0 + \alpha_1 z_{it} + u_i + r_{it} \quad (4.5b)$$

where $\alpha_1 \neq 0$ and Cov(z_{it}, e_{it})=0.

Finally, equations (4.6) are estimated to investigate the effect of scale-up on average cost once controlling for the quality of outreach services.

$$Log(AC_{it}) = \delta_0 + \delta_1 \log Y_{it} + \delta_2 q_{it} + u_i + e_{it} \quad (4.6a)$$

$$Log(Y_{it}) = \alpha_0 + \alpha_1 z_{it} + \alpha_2 q_{it} + u_i + r_{it} \quad (4.6b)$$

where $\alpha_1 \neq 0$ and Cov(z_{it}, e_{it})=0

There are two main types of variables that could potentially affect the average cost only through NGO size. The first type relates to variables determining the revenues received from the funding agency through the state level partner. In Avahan, grants are released every year to state level partners depending on the review of past progress reports, future planned activities and intended coverage. Since there is a bargaining process between the NGO and the state level partner, there is a priori no exogenous determinant of NGO revenue that could be used as an exclusion restriction. Our construction of a plausible instrumental variable then relates to second type of variable, that explains the size of the high-risk population, without influencing costs: the demography of high-risk population. On such possible variable is an unanticipated climatic shock that may impact high risk population size and leads to an exogenous change in the NGO size. Droughts in India are frequent and severe – and it is plausible that they will influence both the demand and supply of sex work. In fact, a drought acts as a negative income shock and is expected to affect negatively the demand for prostitution, which should then in turn have a negative effect on the quantity of sex workers. For those reasons, we explore the presence of a drought in the intervention zone of the NGO as a proxy for a change in the number of high-risk populations. This variable is constructed by using the Standardized Precipitation Index (SPI) that is computed using the data of Schneider et al. (2011). The SPI proposed by McKee, Doesken, and Kleist (1993) is calculated by first, fitting a gamma probability density function to the frequency distribution of rainfall over the reference period, here 1950 to 2000. The probability density function is then used to determine the cumulative probability of a particular precipitation level for a chosen time scale. Finally, the calculation is transformed into a normal distribution with a mean of 0 and a variance of 1 to obtain the SPI. The values of the SPI can be grouped into various classes, where negative values indicate rainfalls below normal, and positive values indicate above normal rainfall. A SPI below -1 indicates dry conditions; the drought variable thus takes the value of 1 when the SPI is below -1 and 0 otherwise. We obtained information on the intervention cities and villages of each NGO between 2004 and 2007. We then computed the SPI for each NGO by considering the average rainfall in its intervention area. The intervention area was defined to be a radius of 45 kilometres around the NGO location and the radius was defined as the average distance of the intervention sites to the NGO location.

Using an instrumental variable can only be justified where it is truly exogenous; it has a strong relationship with cost, and is a valid exclusion restriction i.e. uncorrelated directly with average cost. Firstly, regarding the exogeneity of the instrument, using the drought insures that the relationship between drought and the number of persons reached does not go in both directions. Secondly, regarding the weakness of the instrument, a potential problematic issue is that a drought may not affect high-risk populations in rural and urban areas in the same way. In fact, a drought may decrease the number of sex workers in rural areas through the decrease in demand by agricultural income of farmers. However, a drought may simultaneously increase the number of sex workers in urban centres, assuming that (1) rural women may take up sex work if they consider that it can offer them a higher and more sustainable income than farming activities and (2) sex workers will migrate to places where the demand will be higher. The occurrence of (1) and/or (2) is an issue only if it results in a greater demand for prostitution is urban centres, which may not be the case since droughts may increase the price of non-tradable goods in cities as well, we may then also observe a lower demand for prostitution in urban centres. If NGOs located in urban areas are able to reach more high-risk populations after a drought has occurred then it would result in the weakness of the IV since the negative effect of the IV on the number of persons reached in the rural parts of the NGO intervention may be offset by its positive effect in the urban ones. We therefore present F-stats in the first stage equation to show evidence of a strong relationship between the drought and NGO size. Additionally, the failure of monotonicity is also not likely to be an issue here since we do not expect any heterogeneous effect of scale on the average cost.

Thirdly, regarding the validity of the exclusion restriction, the instrument will be invalid if a drought affects prices of commodities bought by NGOs, NGO labour and transport costs. Regarding commodities prices (mainly condoms and STI kits prices), one may want to note that commodities are not bought by NGOs on the local market but are distributed from state level partners. State level partners purchase in bulk and distribute to NGOs based on intend of their requirement. Commodities prices are negotiated by the state level partner with the supplier and commodities are provided to all NGOs contracting with the state level partner at the same price. Regarding the effect on labour, the instrument would also be invalid if it negatively affects cost through a reduction in labour cost. In fact, NGO staff facing a diminution of their real wage due to the increase in prices generated by a drought may demand higher wages, we would then observed a decreasing effect of drought on labour costs, which is not the case as shown in Appendix 1. Moreover, since labour price is fixed and decided at the beginning of the financial year, a change in wage within the year is not likely to occur. In order to explore this potential pathway, total labour cost was regressed on the NGO size and the drought variable in the panel model with NGOs fixed effects, and the drought variable was not a predictor of labour costs as shown Appendix 1. Note that similar conclusion is found when regressing recurrent, capital costs as well as NGO, state level partner and BMGF costs on the drought variable. Regarding the effect on transport cost, conversely to floods, droughts will have no effect on the quality of roads. Regarding fuel cost, there are not many reasons to believe that they could be affected by a drought especially in

India where fuel prices were regulated, deregulation of petrol occurred in 2010 while diesel and kerosene prices are still under the Government control.

Lastly, if drought was positively affecting average cost through an increase in input prices, this should result in an overestimation of the effect of scale on the persons reached after correcting for the endogeneity, thus using the drought as an instrumental variable ensures that we would never conclude wrongly that there are no economies of scale.

In the sample, 11% of the NGOs have experienced a drought in their intervention area between 2004 and 2007, although this proportion has varied a lot over time since the drought that occurred in 2004 affected 31% of the NGOs intervention area while 2%, 9% and 1% of the intervention areas were affected by a drought in 2005, 2006 and 2007 respectively. Maps presented in appendix 2 provides further evidence of the drought variability in the districts surveyed.

5. Results

NGO average costs excluding programme costs

In column (4.1) of Table 4, we can see that once the NGO has made investment decisions for the financial year, the economies of scale are high because the NGO is deprived of all leeway to minimize cost over the financial year, since most of the costs are considered as fixed in the short run. If an NGO scale of activity increases by 1% over the year, the average cost decreases by 0.56%.

When we try to capture a relationship between average cost and scale in the longer run assuming that input factors become variable we can see that the level economies of scale decreases. In this case if the NGO size increases by 1%, the average cost decreases by 0.19%. When we restrict the sample to the NGOs that have reached full coverage in column (4.3), we can see that the economies of scale are much lower, although the SE increases due to the lower sample size, resulting in the insignificance of the scale variable.

The causal effect of scale-up on average cost is explored by the use of the system GMM in (4.4) and an IV in (4.5) and (4.6). Overidentification tests are used to test whether the excluded instrumental variables are independent of the error terms and can be considered as valid. The tests are conducted by regressing the residuals from the IV regression on all instrumental variables. The non-rejection of the null hypothesis leads to the conclusion that the overidentifying restriction is valid. The p-values of the overidentification tests Sargan and Hansen tests confirm the validity of the lagged values as instruments in the GMM. The F-stats superior to 10 in the first stage equation of the IV suggests that drought is strongly negatively correlated to scale-up. The two methods correcting the endogeneity of scale suggest that the effect of scale-up on average cost is over-estimated in previous models. Once we correct for the reverse causality, scale has no effect on NGO average cost. An increase in scale in 1% reduced the average cost by 0.11% in the GMM system estimate, while IV estimates suggest that there are no economies of scale.

When quality proxies are added in the IV estimate, we can see that providing outreach services of a higher quality increases NGO average cost. Then it is important to question whether scale-up has not occurred at the cost of quality. To explore this possibility, the number of STI visits and the number of condoms distributed per person reached were regressed on NGO size. We find that an increase in scale has no effect on the number of STI visits per person reached and is positively correlated to the number of condoms distributed per person reached. It is conceivable that when NGOs reach full coverage of their estimated population, and with no incentive to increase beyond this they may decide to maximise future budgets by decreasing the intensity of the services provided. However, this was not found in the data when running the estimates only when PR>EP. This suggests that the incentives given by the Bill and Melinda Gates Foundation have allowed NGOs to maximise quantity and that this did not occur at the cost of quality.

Insert Table 4

NGO average cost including programme costs

When focusing on total NGO average costs presented in Table 5, we can see that economies of scale are very high. This is due to the fact that national level support cost is primarily characterised as fixed cost. Our results suggest that while in the very short run, an increase in 1% in NGO size would result in a decrease in the average cost of 0.88%, this percentage will drop to 0.61% in a long run perspective. It is interesting to note that the GMM estimator used to correct for reverse causality leads to similar result than the panel fixed effects estimator, suggesting that the reverse causality issue is not of importance when programme costs are included. This illustrates the fact that NGOs have little room for manoeuvre in determining their total average cost.

Insert Table 5

Predicted and actual costs when scaling-up the initiative are presented in Appendix 3.

6. Discussion

We find that the scale up of HIV prevention in India is associated with high economies of scale using a method that is robust for endogeneity biases; this finding suggests that scaling up HIV prevention services is feasible and is a source of efficiency. Additionally, we find no evidence of diseconomies of scale in our sample of NGOs since only 4 Avahan NGOs out of the 111 included in our analysis have experienced diseconomies of scale. We demonstrate that this finding is implicit to Avahan design and mechanism funding in Avahan since NGOs funding was based on target objectives and there was no incentive to reach more people than what the NGOs got funding for.

When we consider only the cost incurred at the NGO level, we find no evidence of economies of scale once we correct for reverse causality and other potential sources of endogeneity. This latter finding conflicts with previous descriptive evidence, and may be justified in several ways. Firstly, most of the high-risk populations reached (female sex workers and men who have sex with men) are street-based, preventing the peer-educators to reach many individuals at one time. Secondly, NGOs have to provide a mix of services that includes STI treatment and condom distribution preventing from receiving high specialization gains. Finally, most of consumables (condom, STI kits) are bought at a fixed price that does not vary with the size of the NGOs.

But our results nevertheless suggest that for those planning the scale-up of HIV prevention services more generally at the national and global level, cost estimates should include the presence of high economies of scale, driven by large programmatic costs. Our findings suggest that the total cost function would, after a short initial increase; rapidly settle to a steady state. This is broadly in line with the approach taken by the recent resource estimates made by UNAIDS for their Strategic Investment Framework. Our findings also suggest that programmatic costs should not be allocated as total mark-up of site level costs; but should primarily be treated as a fixed costs in any resource requirements model. The question still remains on how fixed programmatic costs vary as one adds on new services; so this only applies on when estimating costs of a set HIV prevention package.

The most important policy recommendation to draw from our results is that cost savings are possible by increasing the size of the NGOs to expand coverage. Consideration to encouraging NGOs to merge or work together may realise further economies of scale. However, any efficiency gains from encouraging large NGOs to provide services must be balanced with other possible benefits from small NGOs. Unlike private for profit organisations, the ability of an NGO to offer services to marginalised populations may lie in its roots and connections with the population it serves. As we observe no diseconomies of scale, we can assume that the Avahan design prevents from reaching the optimal size of NGO. On the other side, avoiding to have diseconomies of scale is attractive and for this reason the Avahan design may be then of interest for African low and middle countries planning to scale up HIV prevention.

Over the period studied NGOs have demonstrated that it is feasible to substantially increase the quantity of persons reached; but that this may have been influenced by information and programmatic incentives. In fact, by looking at the distribution of persons reached per month presented in Appendix 4, it can be observed that in 2005, NGOs that reached the estimated number of targeted population did not put as much effort to reach new populations as NGOs that did not reach the estimated number of targeted high-risk population. This may be explained by the fact that many NGOs entered Avahan in 2005; and these NGOs and may have lacked information regarding the estimated population. However, we can see from Appendix 4 that the scale trend follows the same pattern in 2006 and 2007 for NGOs that did and did not reach the estimated number of targeted population. This indicates that NGOs that have reached estimated population still continue to reach a similar proportion of new population member than the ones that had not reached the estimated population at this time.

The methodological implications of the study are that it may be important for future analyses of economies of scale of health service costs to consider that scale-up is endogenous to average cost. Reverse causality appears to be the main source of endogeneity at the NGO level since the use of the panel estimator with NGO fixed effects allows controlling for time-invariant characteristics of the NGOs that may be correlated with scale. However, we also recommend that those undertaking similar analysis with cross sectional data should give a special attention to the endogeneity issue, since the effect of scale on average cost could be biased due to time-constant unobserved heterogeneity. The direction of the bias in this case will depend on the correlation between those omitted variables and scale. Then, it is a priori hard to know how the scale coefficient will be biased, which justifies the use of appropriate methods to deal with endogeneity. Overall our conclusion that there are no economies of scale when excluding programme costs can be clearly seen to rely heavily on the validity of the instrument variables used.

Our original findings however have to take into account that this study has some limitations. Firstly, our findings at the programme level are somewhat dependent on the methods we used to allocate programme level costs. We are confident that our estimation of programme costs is robust given the work we have conducted to track those costs and for untracked costs to understand the determinants of above level costs. This is mainly explained by the fact that most of the above level costs have been allocated to specific NGOs and could be tracked easily. Although our methods are robust, there remains some uncertainty, since all expenditures were not directly tracked and our results also suggest some measurement error in this regard. Secondly, one may want to note that we developed a parsimonious model in order to better understand the relationship between scale and cost. In fact, the inclusion of many covariates would have complicated the understanding of the biases. For this reason, the study did not aim to provide a broad analysis of the determinants of average cost and this will be conducted in a future study. Thirdly, the findings regarding the effect of scale up on average cost only focuses on provider costs. These may be of interest for decision-makers but do not take into account user costs or other societal factors such as the accessibility of the NGOs. In the sample, most of the NGOs operate alone over a single district, thus increasing NGOs size may also result in other planning issues that cannot be predicted. Fourthly, although we see no inverse relationship between scale and quality, our measures of quality are service orientated, and may not capture the greater complexity of providing acceptable services to high-risk populations. Finally, while we present an analysis of one of the largest datasets available globally on HIV prevention costs, our study remains highly context specific, so we strongly recommend that similar studies are carried out in other settings. These studies need to be planned in advance, as the Avahan data is unique in the sense it was captured prospectively throughout the scale up process. This has enabled the use of panel estimation, thus strengthening our findings.

Conclusion

We quantified the degree of economies for scale that have resulted from the scale-up of the Avahan initiative; the largest HIV prevention project conducted so far. We use GMM and IV approaches to estimate unbiased economies of scale. We find that the scale-up of

Avahan has generated high economies of scale suggesting that cost savings are possible when scaling-up HIV prevention interventions in low and middle income countries.

Table 1: Total economic costs by organisational level and input from 2004 to 2008 (in US$ 2008)

State level Partner

INPUT	2004-05	%	2005-06	%	2006-07	%	2007-08	%	Total	%
Capital cost	3,21,707	10%	7,10,314	9%	7,40,217	8%	8,28,565	9%	26,00,803	9%
Personnel	14,61,108	44%	33,26,119	43%	33,46,931	37%	37,94,869	40%	1,19,29,028	40%
Travel	2,60,931	8%	5,83,292	8%	5,52,527	6%	7,94,457	8%	21,91,207	7%
Building operating & maintenance	1,28,889	4%	6,85,979	9%	8,75,273	10%	8,39,026	9%	25,29,167	9%
Commodities and supplies	3,15,164	9%	9,28,084	12%	11,37,772	13%	11,30,847	12%	35,11,867	12%
Monitoring & Evaluation	4,73,509	14%	5,78,504	7%	7,26,454	8%	4,30,071	5%	22,08,540	7%
Trainings	3,02,135	9%	6,12,627	8%	12,48,216	14%	10,27,480	11%	31,90,457	11%
Indirect Expenses	69,596	2%	3,49,340	4%	3,89,641	4%	6,58,639	7%	14,67,216	5%
Grand Total	33,33,038	100 %	77,74,257	100 %	90,17,032	100 %	95,03,954	100 %	2,96,28,284	100 %

District Level (NGO)

INPUT	2004-05	%	2005-06	%	2006-07	%	2007-08	%	Total	%
Capital cost	3,35,362	15%	7,71,906	11%	9,86,912	9%	12,42,946	9%	33,37,127	10%
Personnel	9,88,547	43%	32,48,881	47%	45,57,267	43%	63,35,755	44%	1,51,30,451	44%
Travel	1,48,326	6%	4,56,460	7%	6,96,232	7%	9,73,823	7%	22,74,841	7%
Building operating & maintenance	1,61,702	7%	3,86,134	6%	5,15,022	5%	11,09,663	8%	21,72,521	6%
Commodities & Supplies*	4,30,133	19%	17,24,818	25%	32,64,794	31%	39,38,449	27%	93,58,194	27%
Monitoring & Evaluation	1,19,348	5%	91,711	1%	89,520	1%	1,52,707	1%	4,53,286	1%
Trainings	1,03,761	5%	2,28,316	3%	3,88,067	4%	6,47,648	4%	13,67,792	4%
Indirect Expenses	7,958	0%	33,313	0%	1,51,883	1%	1,40,755	1%	3,33,906	1%
Grand Total	22,95,137	100	69,41,539	100	1,06,49,697	100	1,45,41,746	100	3,44,28,119	100

*drugs, condoms and IEC materials

Table 2: Average cost per person reached over time in US$2008

Year	NGO average cost	National level cost		Total average cost
		State Level Partner cost	BMGF level and Pan-Avahan capacity building partners cost	
2004	65	180	211	456
2005	64	80	137	281
2006	59	54	58	169
2007	61	40	38	139
Total	*62*	*75*	*94*	*231*

Table 3: Descriptive statistics of variables in the sample

Variable	Obs	Mean	Std. Dev.	Min	Max
AC_tot	387	230.884	451.1441	36.8	5597.4
Log(AC_tot)	389	5.060426	0.803372	3.605498	10.37525
AC_ngo	386	61.98771	56.46357	9.259	440.8
Log(AC_ngo)	388	3.905978	0.729109	2.230014	8.261216
Y	388	1868.66	1729.984	20	12071
Log(Y)	388	7.147942	0.972623	2.995732	9.398561
STI/Y	388	0.772854	0.81042	0	10.75
Condom/Y	388	248.4381	242.5125	0	1561.993

Table 4: Relationship between scale and NGO average cost

VARIABLES	(4.1) LogAC_ngo$_{it}$	(4.2) LogAC_ngo$_{it}$	(4.3) LogAC_ngo$_{it}$	(4.4) LogAC_ngo$_{it}$	(4.5a) LogAC_ngo$_{it}$	(4.5b) $logY_{it}$	(4.6a) LogAC_ngo$_{it}$	(4.6b) $logY_{it}$
$logY_{it}$	-0.558***	-0.193***	-0.0679	-0.115	-0.0176		0.0716	
	(0.0676)	(0.0546)	(0.0848)	(0.0680)	(0.207)		(0.187)	
2005 (ref:2004)	0.636***							
	(0.0839)							
2006	0.851***							
	(0.101)							
2007	1.061***							
	(0.125)							
STI/Y_{it}							0.262***	-0.298***
							(0.0685)	(0.0558)
Condom/Y_{it}							0.000591**	0.000134***
							(0.000291)	(0.000201)
Drought						-0.6097***		-0.4236***
						(0.151)		(0.145)
Constant	7.138***	5.261***	4.198***	4.694***	4.214***	10.418***	3.436***	10.523***
	(0.421)	(0.390)	(0.647)	(0.488)	(0.893)	(0.562)	(1.004)	(0.649)
NGO Fixed Effects	YES	YES	YES	YES	YES	YES	YES	YES
Observations	386	386	185	386	367		367	
R-squared	0.451	0.105	0.010		0.045		0.368	
Number of NGOs	130	130	90	130	111		111	
Sargan test p-value				0.84				
Hansen test p-value				0.61				
Cragg-Donald F-stat						16.25		9.10

Standard errors are in parentheses, ** p<0.01, ** p<0.05, * p<0.1. Column (1) presents the results considering the very short run. Column (2) is the basic estimation with NGO fixed effects. Column (3) is the same model as in column (2) estimated for the sub-sample of NGOs that have reached a higher number of high-risk populations than the one estimated in the catchment area. Column (4) is the GMM model that uses lagged values of the scale as instruments for scale. Column (5a) is the IV estimation that uses drought as an instrument. First stage equation for this model is presented in column (5b). Column (6) is the same model than in Column (5a) except that it includes the quality of outreach as covariate. First stage equation for this model is presented in column (6b)

Table 5: Relationship between scale and average cost at the programme level

VARIABLES	(4.1) LogAC_tot$_{it}$	(4.2) LogAC_tot$_{it}$	(4.3) LogAC_tot$_{it}$	(4.4) LogAC_tot$_{it}$	(4.5a) LogAC_tot$_{it}$	(4.5b) $logY_{it}$	(4.6a) LogAC_tot$_{it}$	(4.6b) $logY_{it}$
$logY_{it}$	-0.888***	-0.685***	-0.614***	-0.670***	-0.767***		-0.823***	
	(0.0260)	(0.0224)	(0.0344)	(0.0494)	(0.0938)		(0.116)	
2005 (ref:2004)	0.496***							
	(0.0429)							
2006	0.543***							
	(0.0494)							
2007	0.621***							
	(0.0543)							
STI/Y_{it}							0.0347	-0.298***
							(0.0426)	(0.0558)
Condom/Y_{it}							0.000745***	0.000134***
							(0.000182)	(0.000201)
Drought						-0.6097***		-0.4232***
						(0.151)		(0.145)
Constant	10.91***	9.931***	9.328***	9.838***		10.418***		10.523***
	(0.163)	(0.161)	(0.263)	(0.354)		(0.562)		(0.649)
NGO Fixed Effects	YES	YES	YES	YES	YES	YES	YES	YES
Observations	387	387	186	387	368		368	
R-squared	0.867	0.786	0.770		0.774		0.839	
Number of NGOs	130	130	90	130	111		111	
Sargan test p-value				0.32				
Hansen test p-value				0.04				
Cragg-Donald F-stat						16.23		9.07

Standard errors are in parentheses, ** p<0.01, ** p<0.05, * p<0.1. Column (1) presents the results considering the very short run. Column (2) is the basic estimation with NGO fixed effects. Column (3) is the same model as in column (2) estimated for the sub-sample of NGOs that have reached a higher number of high-risk populations than the one estimated in the catchment area. Column (4) is the GMM model that uses lagged values of the scale as instruments for scale. Column (5a) is the IV estimation that uses drought as an instrument. First stage equation for this model is presented in column (5b). Column (6) is the same model than in Column (5a) except that it includes the quality of outreach as covariate. First stage equation for this model is presented in column (6b)

Figure 1: Relationship between average cost and scale-up

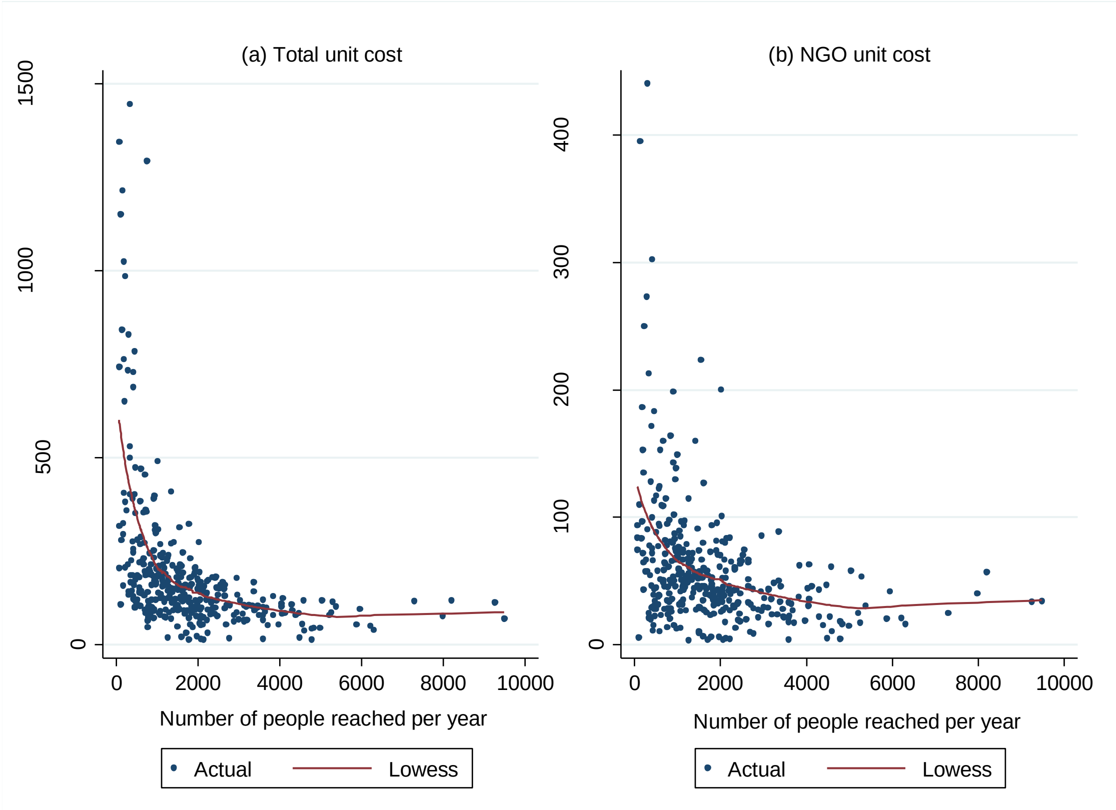
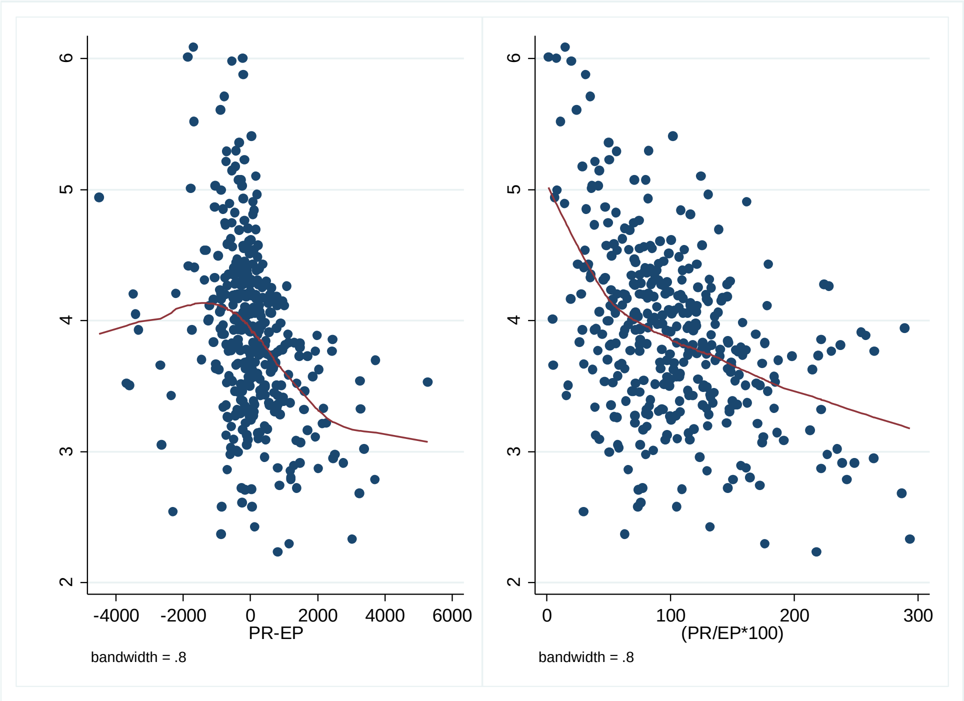


Figure 2: Relationship between coverage and average cost



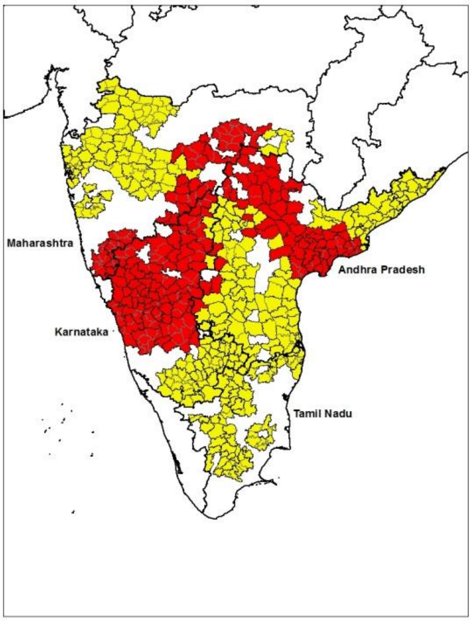
Appendix 1: Effect of drought on several costs

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Log(NGO_cost)	Log(SLP_cost)	Log(BMGF_cost)	Log(recurrent_cost)	Log(capital_cost)	Log(labour_cost)
$logY_{it}$	0.794***	0.187***	0.0686*	0.936***	0.466***	0.766***
	(0.0545)	(0.0347)	(0.0372)	(0.0610)	(0.0433)	(0.0572)
Drought	-0.123	0.0856	0.0621	-0.147	-0.0410	-0.140
	(0.100)	(0.0633)	(0.0822)	(0.116)	(0.0820)	(0.103)
Observations	367	368	368	367	367	367
R-squared	0.676	0.171	0.025	0.643	0.561	0.624
Number of ngo	111	111	111	111	111	111

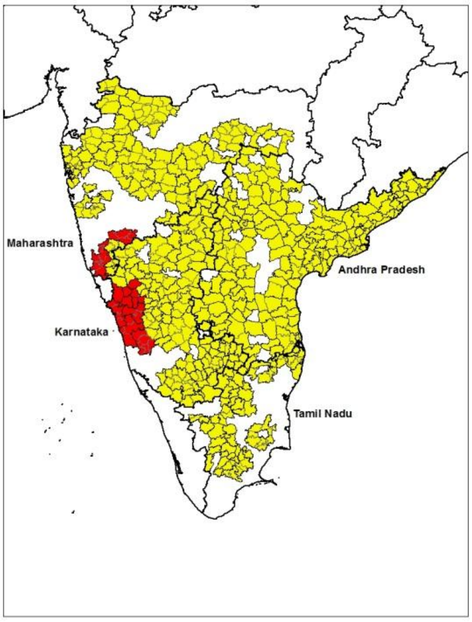
Robust standard errors in parentheses, *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Appendix 2: Location of droughts

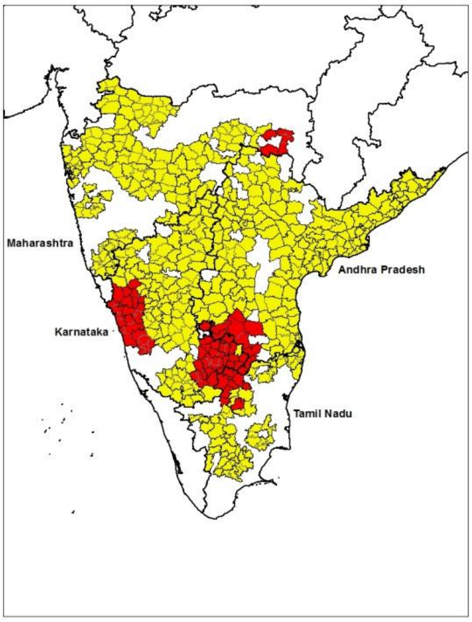
2004



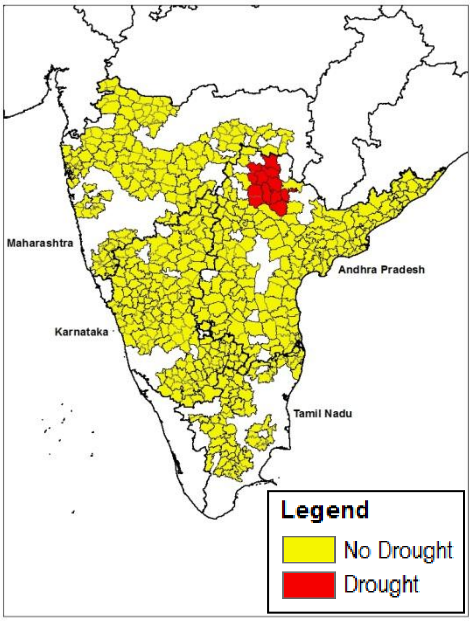
2005



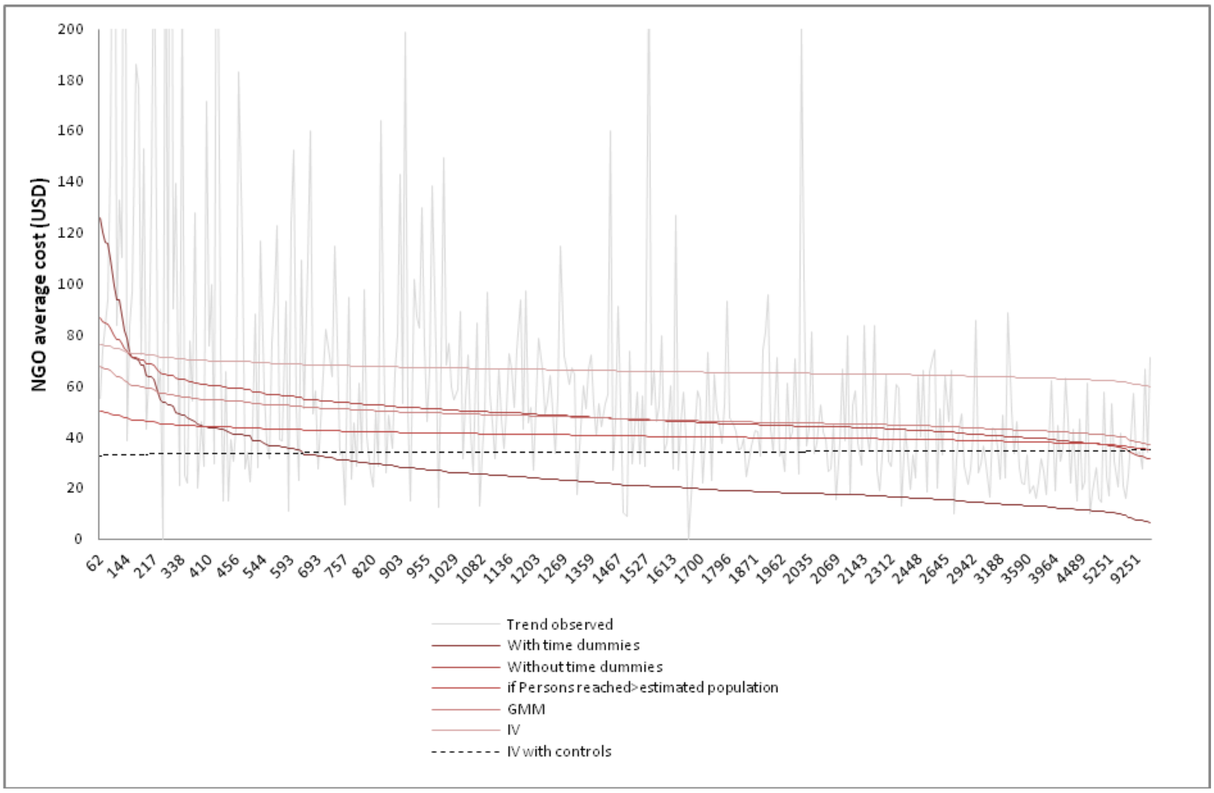
2006

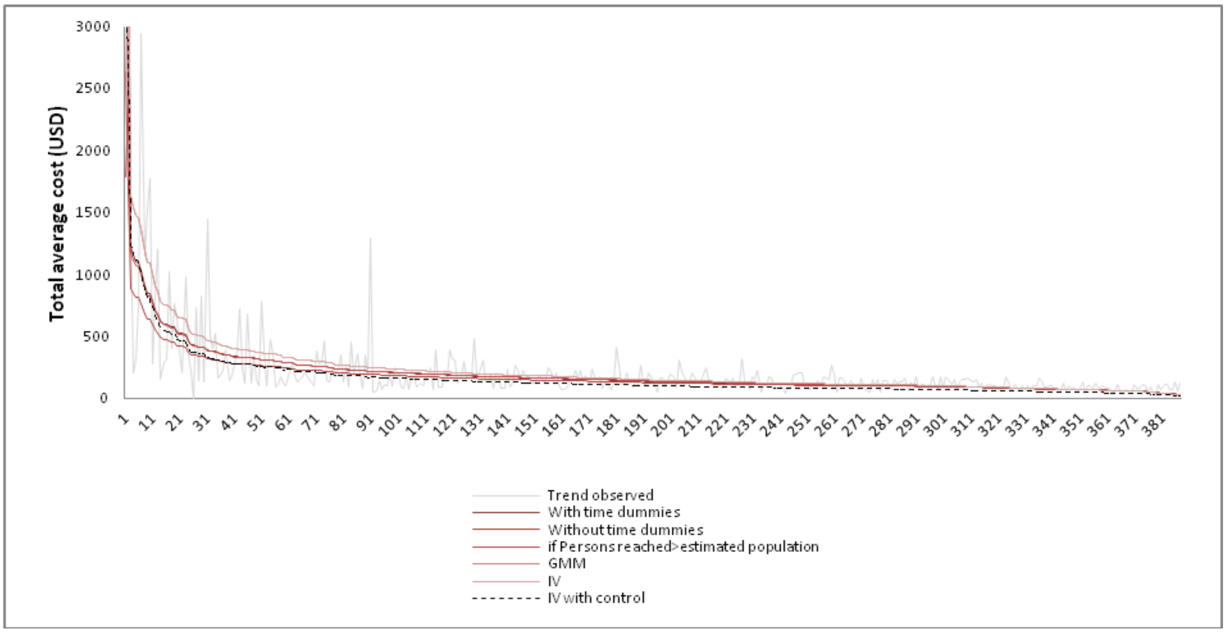


2007

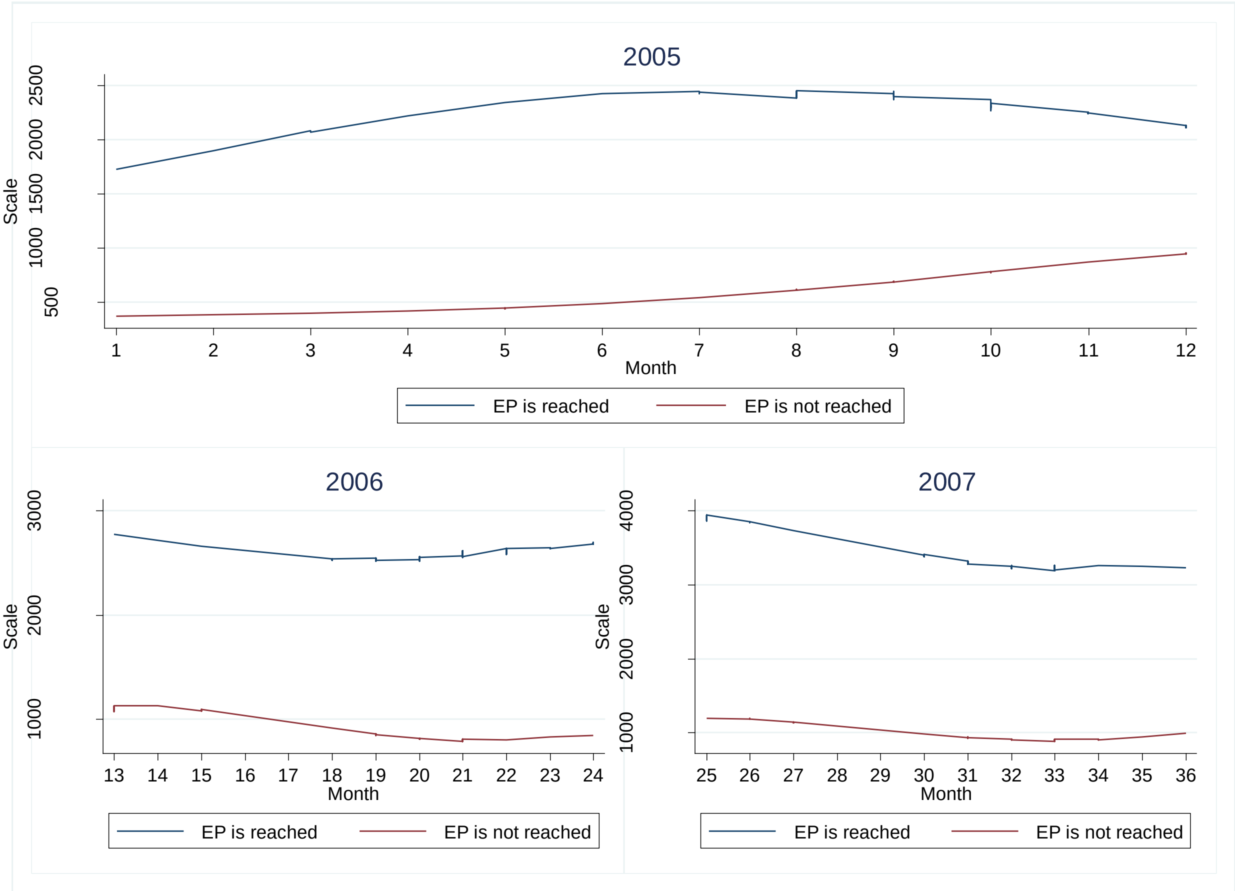


Appendix 3: Predicted and actual effect of scale-up on average cost





Appendix 4: Average trend in scale per year if the estimated population (EP) is reached or not



References

Bill and Melinda Gates Fundation. (2008). Avahan - The India AIDS Initiative: Business of HIV Prevention at Scale. New Delhi, India.

Blanchard, J. F., Bhattacharjee, P., Kumaran, S., Ramesh, B. M., Kumar, N. S., Washington, R. G., & Moses, S. (2008). Concepts and strategies for scaling up focused prevention for sex workers in India. *Sexually Transmitted Infections, 84*(2).

Blundell, R., & Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of econometrics, 87*(1), 115-143.

Chandrashekar, S., Guinness, L., Pickles, M., Shetty, G., Alary , M., Vickerman, P., . . . Vassall, A. (2014). *The costs of scaling up HIV prevention high risk groups: lessons learned from the Avahan Programme in India.*

Chandrashekar, S., Guinness, L., Kumaranayake, L., Reddy, B., Govindraj, Y., Vickerman, P., & Alary, M. (2010). The effects of scale on the costs of targeted HIV prevention interventions among female and male sex workers, men who have sex with men and transgenders in India. *Sexually Transmitted Infections, 86*(1).

Dandona, L., Sisodia, P., Kumar, S. G. P., Ramesh, Y., Kumar, A. A., Rao, M. C., . . . Kahn, J. G. (2005). HIV prevention programmes for female sex workers in Andhra Pradesh, India: outputs, cost and efficiency. *BMC Public Health, 5*(1), 98.

Guinness, L., Kumaranayake, L., & Hanson, K. (2007). A cost function for HIV prevention services: is there a 'u' - shape? *Cost Effectiveness and Resource Allocation, 5*(1), 13.

Guinness, L., Kumaranayake, L., Rajaraman, B., Sankaranarayanan, G., Vannela, G., Raghupathi, P., & George, A. (2005). Does scale matter? The costs of HIV-prevention interventions for commercial sex workers in India. *Bulletin of the World Health Organization, 83*(10), 747-755.

Kumaranayake, L., & Watts, C. (2005). HIV/AIDS Prevention and Care Interventions in Sub-Saharan Africa: An Econometric Analysis of the Costs of Scaling-Up. *South African Journal of Economics, 68*(5), 455-467.

Lave, J. R., & Lave, L. B. (1970). Hospital cost functions. *The American Economic Review*, 379-395.

Marseille, E., Dandona, L., Marshall, N., Gaist, P., Bautista-Arredondo, S., Rollins, B., . . . Lioznov, D. (2007). HIV prevention costs and program scale: data from the PANCEA project in five low and middle-income countries. *BMC Health Services Research, 7*(1), 108.

Marseille, E., Giganti, M. J., Mwango, A., Chisembele-Taylor, A., Mulenga, L., Over, M., . . . Stringer, J. S. A. (2012). Taking ART to Scale: Determinants of the Cost and Cost-Effectiveness of Antiretroviral Therapy in 45 Clinical Sites in Zambia. *PloS one, 7*(12).

McKee, T. B., Doesken, N. J., & Kleist, J. (1993). *The relationship of drought frequency and duration to time scales.* Paper presented at the Proceedings of the 8th Conference on Applied Climatology.

Menzies, N. A., Berruti, A. A., & Blandford, J. M. (2012). The Determinants of HIV Treatment Costs in Resource Limited Settings. *PloS one, 7*(11).

National AIDS Research Institute. (2005-2010). Integrated Behavioral and Biological Assessment Rounds 1&2.

Pickles, M., Boily, M.-C., Vickerman, P., Lowndes, C. M., Moses, S., Blanchard, J. F., . . . Washington, R. (2013). Assessment of the population-level effectiveness of the Avahan HIV-prevention programme in South India: a preplanned, causal-pathway-based modelling analysis. *The Lancet Global Health, 1*(5), 289-299.

Ramesh, B. M., Moses, S., Washington, R., Isac, S., Mohapatra, B., Mahagaonkar, S. B., . . . Subramanian, T. (2008). Determinants of HIV prevalence among female sex workers in four south Indian states: analysis of cross-sectional surveys in twenty-three districts. *Aids, 22*(5), 35-44.

Roodman, D. (2006). How to do xtabond2: An introduction to difference and system GMM in Stata. *Center for Global Development working paper* (103).

Rosen, S., Long, L., & Sanne, I. (2008). The outcomes and outpatient costs of different models of antiretroviral treatment delivery in South Africa. *Tropical Medicine & International Health, 13*(8), 1005-1015.

Schneider, U., Becker, A., Finger, P., Meyer-Christoffer, A., Rudolf, B., & Ziese, M. (2011). GPCC Full Data Reanalysis Version 6.0 at 0.5°: Monthly Land-Surface Precipitation from Rain-Gauges built on GTS-based and Historic Data.

Schwartländer, B., Stover, J., Hallett, T., Atun, R., Avila, C., Gouws, E., . . . Padian, N. (2011). Towards an improved investment approach for an effective response to HIV/AIDS. *The Lancet, 377*(9782), 2031-2041.

Vassall, A., Pickles, M., Chandrashekar, S., Boily, M., Shetty, G., Guinness, L., . . . Vickerman, P. (2014). The cost-effectiveness of HIV prevention for high-risk groups at scale: an economic evaluation of the Avahan programme in South India. *The Lancet Global Health, In press*.

Verma, R., Shekhar, A., Khobragade, S., Adhikary, R., George, B., Ramesh, B. M., . . . Bohidar, N. (2010). Scale-up and coverage of Avahan: a large-scale HIV-prevention programme among female sex workers and men who have sex with men in four Indian states. *Sexually Transmitted Infections, 86*(1).

Wheeler, T., Kiran, U., Dallabetta, G., Jayaram, M., Chandrasekaran, P., Tangri, A., . . . Alexander, A. (2012). Learning about scale, measurement and community mobilisation: reflections on the implementation of the Avahan HIV/AIDS initiative in India. *Journal of Epidemiology and Community Health, 66*(2).

Windmeijer, F. (2005). A finite sample correction for the variance of linear efficient two-step GMM estimators. *Journal of econometrics, 126*(1), 25-51.

*Ethics approval/Statement EA not required

पी.ए.बी.एक्स./PABX : 26588980, 26588707, 26589336, 26589745, 26589873, 26589414
फैक्स/FAX : 011-26588662, 011-26589791, 011-26589258

तार / GRAM : विज्ञानी / SCIENTIFIC
Web-site : www.icmr.nic.in
E-mail : icmrhqds@sansad.nic.in

भारतीय आयुर्विज्ञान अनुसंधान परिषद
INDIAN COUNCIL OF MEDICAL RESEARCH

वी. रामलिंगस्वामी भवन, अन्सारी नगर, पोस्ट बॉक्स 4911, नई दिल्ली - 110 029
V. RAMALINGASWAMI BHAWAN, ANSARI NAGAR, POST BOX 4911, NEW DELHI - 110 029

No. HIV/51/102/2004-ECD-II

Dated: 29/11/05

To,

Dr. Reynold Washington,
Associate Professor,
Institute of Population Health,
St. Jhon's Medical College,
Banglore – 560 034.

Subject: - Monitoring and evaluation of the Avahan project in India: impact assessment and cost-effectiveness analyses using enhanced surveillance methods and mathematical modeling of HIV transmission dynamics.

Sir,

I am pleased to intimate that above proposal has been approved by Secretary, Ministry of Health & Family Welfare, New Delhi from security / sensitivity angle on behalf of the High Level Committee of Secretaries.

Yours faithfully,

[redacted]

(Dr. Deepali Mukherji)
Chief, Division of ECD
For Director General