Welfare and Common Property Rights Forestry: Evidence from Ethiopian Villages[*]

Dambala Gelo[†] and Steven F Koch[‡]

Abstract

Welfare effects associated with a unique multifaceted common-property forestry program in Ethiopia were examined. The combined program impact is estimated via matching, selection models and instrumental variable (IV) methods. Data for the analysis is taken from selected villages in Gimbo district, in southwestern Ethiopia. The program was found to raise the welfare of the average program participant households by approximately 400 Ethiopian Birr (ETB), or 22.5%, and that result was robust to various specifications. Overall, the results suggest that JFM, combined with improved non-timber forest product market linkages, offers one avenue for both rural development and environmental improvement.

J.E.L. Classification: Q23, Q28

Keywords: community forestry, treatment effects, IV, matching and Ethiopia

---

[*] We are grateful to SIDA for financing the data collection of this project through EfD (EDRI/EEPFE), and to *Economic Research Southern Africa* for additional support of this research. We are also grateful for comments received from Madhu Kanna, Abebe Damte, Chitalu Chama and Gauthier Tshiswaka-Kashalala. Their suggestions have greatly improved the manuscript. The usual declaimer applies.

[†] PhD, Department of Economics, University of Pretoria, Private Bag X20, Hatfield 0028, Republic of South Africa, daatu24@yahoo.com.

[‡] Professor and HOD, Department of Economics, University of Pretoria, Private Bag X20, Hatfield 0028, Republic of South Africa, steve.koch@up.ac.za, (O) 27-12-420-5285.

1. INTRODUCTION

The devolution of natural forest management to local communities has become more widespread, due to a growing recognition that local communities are likely to manage forest resources better than the state (Murty, 1994; Agrawal and Gibbon, 1999 and Gauld, 2002). Decentralization is also seen as a means of developing and upholding democratisation, allowing people to engage in their own affairs (Agrawal and Ostrom, 2001). However, improving the management of forests and upholding democracy, is likely to hinge on the ability of forest management decentralization to improve the standard of living of those who are dependent on forests; Angelsen and Wunder (2003) and Sunderlin et al. (2005), amongst others, believe forest management decentralization can reduce poverty.

Management decentralization, within the context of tropical forest management, has taken two major forms: Community Based Forest Management (CBFM) and Joint Forest Management (JFM). CBFM involves a broad set of property rights transfer (exclusion, withdrawal and management), wherein local communities become decision-maker over and the residual claimant on forest products. Given this local nature, it is more commonly adopted where forests don't generate significant national or global externalities, although a general improvement in forest cover remains an objective of these reforms (Robinson and Lokina, 2012). Under JFM, on the other hand, forest ownership remains with the state, while local communities are granted user rights. In other words, exclusion and alienation rights remain with the state, while access, withdrawal and management rights are bestowed on forest user collectives, which enjoy variable access to different forest products and services.

The present study focuses on one JFM program in Ethiopia, described in detail in Section 2. JFM aims at forest conservation by placing significant restrictions on forest timber harvest, charcoaling and agricultural encroachment. Those practices were shown to lead to deforestation under the preceding open access regime. Under JFM, ownership rents associated with improved forest stocks do not accrue to the local communities; hence, there is a need to draw on alternative mechanisms to incentivize forest protection. One possibility is to confer common property right usufruct for non-timber forest products (NTFP), the harvest of which is environmentally less detrimental than timber harvesting or other forest uses (Arnold and Perez, 2001 and Wunder 2001). NTFP usufruct, alone, is not likely to be enough, unless it can yield price premiums for the gatherers of NTFPs, due to the risks and high transaction costs associated with the marketing of these goods (Shillington, 2002, and Neumann and

Hirsch, 2000). In that sense, the public purse, as well as donor incentives, can be used to tackle these impediments, through technical, institutional and infrastructural interventions in the marketing of NTFPs that would help JFM communities benefit from the forest (World Development Report, 2008).

To some extent, this hypothesis, conservation by commercialization (Evans, 1993), has been influential, Arnold (1998) and Wunder (2001).[1] However, evidence of its welfare benefit is scant.[2] In the Indian subcontinent, Neumann and Hirsch (2000), find that the success of state interventions to reduce NTFP marketing impediments is generally not positive, due to corruption, bureaucratic hurdles, inappropriate price setting and rent seeking. Jumbe and Angelsen's (2006) Malawian study, based on propensity score matching and Heckman (1979) selection models, finds a JFM program with contrasting impacts across villages. Likewise, in Nepal, Cooper's (2007) CGE analysis of JFM programs uncovers a welfare loss for all concerned, although outcomes for the poor are worse. Cooper's (2008) panel analysis, though, is more supportive of JFM, finding increased growth in per capita consumption, despite increased inequality. For the most part, the preceding studies consider JFM programs involving local community forest protection in exchange for benefits that could arise from long-term sustainable management – access to fuel wood and non-timber forest products (NTFP) – for own consumption. The Ethiopian program we consider, on the other hand, includes both access rights for own consumption and the possibility for increased market returns for, especially, NTFPs.

Although one criticism of the following study is its inability to separate the program effect of own consumption access rights from that of market linkages, the program examined does include both components, and, therefore, in the light of limited evidence that own consumption access rights offer positive welfare gains, it is worthwhile considering whether or not a combined program could be effective. Unfortunately, separating the effect of the two program components is not possible in this study, as requisite data for a program without marketing incentives is not available. However, considering the significant harvest restrictions on timber products (logging,

[1] Furthermore, the forestry literature contains ample evidence that community forestry is beneficial for the environment. Community forestry is positively associated with increased forest stocks (Nagendra, 2002, Bekele and Bekele, 2005, Kassa et al., 2009, Blomley et al., 2008, and Gobeze et al., 2009), as well as improved biodiversity and carbon sequestration (Klooster and Masera, 2000). Moreover, Edmonds (2002), Yadev et al. (2003) and Bluffstone (2008) report reduced forest resource extraction efforts by program households, implying increases in the forest stock.

[2] The World Bank (2008) has suggested one avenue for its efficacy: increases in the forest stock may increase the return to other natural and human assets. Improved forest stocks could protect the quantity and quality of water, favorably impacting household health and labour productivity. Increased forest cover may also help control soil erosion and flooding, resulting in an increase in land productivity. Similarly, increased forest stocks potentially reduce timber product and non-timber product collection times, unleashing labor for other purposes.

charcoaling) and on agricultural encroachment associated with this program, compared to the previous open access regime, along with the potential for NTFP marketing to increase the proportion of the final product price to be returned to farmers, we are led to believe that the market incentive effect (price effect) dominates the forest restriction effects (quantity effect). As noted above, providing incentives to farmers to eschew short-term gains in favour of medium- to long-term payoffs are likely to be important determinants in the success of such programs, and, therefore, should not be ignored, when considering the multifaceted goals of afforestation and rural development.

It is the aforementioned multifaceted objective that motivated the present study. Therefore, this study aims to evaluate the welfare impact of a JFM program augmented by the provision of market-based incentives, through NTFP marketing. For the analysis, a combination of econometric methods was used, including propensity score matching and IV methods via 2SLS. To examine the robustness of the estimates, bounds methods associated with matching and polychotomous choice selection models, as well as various instrument identification tests, related to IV, were also considered. We applied these methods to data collected from households living close to forests in selected villages of the Gimbo district, in southwestern Ethiopia. To summarize, the analysis points to both economically and statistically significant increases in welfare, which can be attributed to participation in this JFM program.

Therefore, this study contributes by adding to the small, but growing, literature related to the evaluation of environmental policies in developing and emerging countries, while providing evidence of the effect of decentralized forestry management programs that are augmented by market-based incentives, through the marketing of NTFPs. Given the widespread devolution of natural forest management throughout developing and emerging economies (Agrawal and Ostrom, 2001, Bluffstone, 2008), which is invariably based on theoretical predictions, as well as anecdotal evidence from local case studies, rigorous empirical analysis of the welfare impact is needed to inform such policies. Our results provide support for the hypothesis that decentralized forestry management, combined with a complementary market access policy, has the potential to raise the welfare of program participants, and that result is robust to a range specifications.

The remainder of the paper is organized as follows. Section 2 discusses the background to common property forestry in Ethiopia, as well as the context of the study. Section 3 describes the data collection efforts, while Section 4 discusses the econometric framework that informed the empirical strategies. Section 5 presents results and discusses those results. Finally, Section 6 concludes the analysis.

2. COMMON PROPERTY FOREST MANAGEMENT IN SOUTHWEST ETHIOPIA

Much like citizens of a number of developing and emerging economies, Ethiopians depend heavily on forest resources, and the reasons for that dependence are many. Ethiopia's modern energy sector is not well developed, such that biomass fuel consumption incorporates 96% of total energy consumption (Mekonnen, 1999, Mekonnen and Bluffstone, 2008), 82% of which comes from fuel wood (World Bank, 1994). Given the lack of modern energy development, Mekonnen and Bluffstone (2008) expect this dependency to continue, and, most likely, to grow. In addition to providing fuel for energy, the forest offers agricultural risk mitigation services, providing alternative sources of income (Delacote, 2007).

In recognition of the importance of forest resources and the realization that deforestation rates, currently at 8% (World Bank, 2005), are not likely to decrease soon, Ethiopia has begun to implement a new set of forest policies (Mekonnen and Bluffstone, 2008). One of those is the decentralization of forest management to the communities located near those resources. Due to that policy, a number of programs have been implemented in Chilimo, Bonga, Borena and Adaba Dodola (Neumann, 2008, and Jirane et al., 2008), with the general objectives of arresting deforestation, while improving the welfare of those largely dependent on the forest. Although the 2007 Ethiopian forestry policy supports decentralization (Mekonnen and Bluffstone, 2008, and Nune, 2008), bilateral donors, such as the GTZ and JICA, as well as NGOs, including Farm Africa/SOS-Sahel, are also supporting these programs.[3] These external actors have provide financial support and helped mediate between the local communities and the local and regional governments. In Bonga, which is the site of this analysis, Farm Africa/SOS-Sahel supported the implementation of participatory forestry management (PFM); more than six PFM programs have been established to improve the

[3] The Deutsche Gesellschaft für Technische Zusammenarbeit (German Technical Cooperation), GTZ, is a bilateral agency mainly engaged in urban and rural development and environmental protection endeavors in Ethiopia. The Japan International Cooperation Agency, JICA, provides technical cooperation and other forms of aid that promote economic and social development. Farm Africa is a UK based registered charity, which operates mostly in eastern Africa, focusing on agricultural development and, to some extent, on natural resource management. SOS-Sahel is also a UK based registered charity focusing, primarily on operations in Africa's arid regions, such as the Sahel.

management of about 80,066 ha of natural forest (Jirane et al., 2008).[4]

As might be expected, donor involvement hinges, in part, on whether or not the donor believes the program will be successful. Therefore, Farm Africa/SOS-Sahel set intervention preconditions focusing on the possibility of success. Effectively, the level of local community and government concern over the current forest situation and the donor's perception of the degree of forest exploitation are important components of these preconditions. Once a forest unit has been provisionally accepted, the location of the forest is topographically identified and then demarcated in the field. Within the provisionally identified forest unit, information related to available forest resources was required, as was information related to past and present management practices. Finally, it was necessary to develop an understanding of prevailing forest management problems, forest uses and forest user needs (Lemenih and Bekele, 2008).

A number of observations emerged from this multi-step process. Importantly, agricultural encroachment into forests, illegal logging, and the harvest of fuel wood, for either direct sale or charcoal production, stood out as major deforestation threats, and these activities were most often associated with unemployed urbanites and a heavy concentration of individuals from the Menja tribe.[5] These observations led Farm Africa/SOS-Sahel and local government to target PFM interventions towards forests surrounded by significant Menja populations (Lemenih and Bekele, 2008, Bekele and Bekele, 2005). Although the Menja population was the overriding eligibility criterion, other criteria, including the degree of agricultural encroachment, population pressure, the forest's status, and the forest's potential to produce non-timber forest products, were considered to a varying degree.

Once intervention sites had been identified, Farm Africa/SOS-Sahel began negotiations and discussions with all stakeholders. However, since skepticism regarding PFM was rife within both the local government and the local communities, Farm Africa/SOS-Sahel provided PFM training for all stakeholders (Bekele and Bekele, 2005). In addition to problems related to

---

[4] PFM formation has undergone a series of steps. Those steps include: identifying forest units to be allocated to forest user groups (FUGs); defining forest boundaries, through government and community consensus; and facilitating the election of PFM management teams (Neumann, 2008; Jirane et al., 2008 and Bekele and Bekele, 2005).

[5] The Menja tribe in Bonga province is a minority ethnic group that is entirely dependent on forests for their livelihood. They are generally ostracized, and commonly referred to as fuel wood sellers (Lemenih and Bekele, 2008, Gobeze et al., 2009, and Bekele and Bekele, 2005).

skepticism, negotiations with regard to PFM participation and PFM forest boundaries were fraught with difficulties. Whereas PFM membership is meant to include those who actually use a particular area of the forest – regardless of their settlement configuration, clan and/or ethnicity – membership negotiations involved both collective and individual decisions. The result was that the entire community determined eligibility based on customary rights, as well as the existing forest-people relationship, which includes the settlement of forest-users, the area of forest-use, and whether or not forest-use was primary or secondary (Lemenih and Bekele, 2005).[6] Program participation amongst eligible households, however, remained voluntary, as long as the household satisfies the eligibility criterion and abides by the PFM's operational rules. Eligible households that choose to participate form Forest User Groups (FUG). Those choosing not to participate must revert to using the nearest non-PFM forest, which, in effect, is a forest that operates under the *status quo*; that forest is unregulated, and access is open to all. It is assumed that household participation is determined by the perceived costs and benefits of the PFM, a perception that is likely affected by training and other household-specific circumstances, which is driven, in large part, by program eligibility.

Experts from Farm Africa/SOS-Sahel and local governments, in collaboration with FUG Members, develop Forest Management Plans (FMP) stipulating the rights and duties of its members.[7] Commonly, each member is required to participate in forest development (planting new trees for the enrichment of the existing forest), guard against fire, vandalism (including unauthorized tree cutting) and agricultural encroachment (clearing forest for agricultural land acquisition). In return, each individual member enjoys two kinds of rights over forest products: (1) a private right and (2) a collective right. Privately, the forest can be used for livestock production, collecting wood for private use (including energy and farm implement construction), harvesting medicinal plants for own consumption, and beekeeping, all subject to management committee approval. The harvest of timber, forest coffee, and spices is a collective right, belonging to the Forest User Cooperative (FUC), leading to benefits that must be distributed across the membership, although 30% of total income is retained by the FUC (Bekele and Bekele, 2005, Lemenih and Bekele, 2008).[8] The FUC operates as a clearing house; members harvest and deliver products to the FUC, which sells the products on both national

[6] Primary users are those who use the forest more frequently, permanently or directly, whereas secondary users are those using the forest less frequently and those who are located farther from the forest boundary (Lemenih and Bekele, 2008).
[7] An elected management committee, comprising of a chairperson, a deputy chairperson, a secretary, a cashier and an additional member, implements the operational plans.
[8] Forest User Cooperatives are fully implemented and operational forest user groups (Jirane, 2008).

and international markets, disbursing the proceeds as a membership dividend.

Possibly the most important aspect of the program is that NGOs, along with the regional government, provide FUCs with assistance in marketing, processing, grading, certification, storing, and market access. Although there is no doubt that the additional assistance confounds the program's effects, the assistance is best viewed as a subsidy to engender participation in the program. Previous research has shown that forest cover and forest productivity have improved under the PFM (Bekele and Bekele, 2005 and Limineh and Bekele, 2008). However, that improvement, on its own, is not likely to offset either the participation cost – the immediate sacrifice of free forest access in the newly established program forest – or the costs of harvest restrictions, imposed by the program, that are generally necessary for the long-term revitalization of forest resources. To offset both the upfront participation costs and the long-term harvest restriction costs, a subsidy of this nature may be necessary. NTFP marketing and market linkage assistance for FUCs offers a potential subsidy. Although forest coffee, for example, sells for as much as ETB 60/kg in non-PFM regions, Shumeta et al., (2012) find that revenues from the sale of forest coffee are not equally split; approximately 13% goes to farmers, while 87% goes to intermediaries in the supply chain. Moreover, the same study argues that intermediary average profit is as much as 40 times higher than that of farmers. If the PFM program can capture an additional proportion of the sales price, and pass that on to the farmer, farmers can benefit from participation, as can the environment. If those gains cannot be realized, farmers will either not subscribe to the program or the program will suffer from attrition; regardless, the environment would be expected to continue to suffer. Therefore, although the program is multifaceted, with potentially confounding effects, the subsidy is likely to be crucial to acceptance by the local residents.

3. METHODOLOGY

Program impact is defined as the difference between the observed outcome and the counterfactual outcome – the outcome that would have obtained had the program not been taken-up (Rubin 1973; Heckman et al., 1998 and Cobb-Clark and Crossley, 2003). As is well understood in the program evaluation literature, counterfactuals are unobservable; at any point in time an individual is either in one state or the other. In this study, a quasi-experimental approach to identifying the appropriate counterfactual is followed, accepting that program participation is not random. As such, appropriately controlling for participation decisions is tantamount to identifying the program impact.

The theoretical foundations of the analysis follow Roy (1951). Accordingly, farmers choose whether or not to participate, and that decision is assumed to depend on the farmer's expectation of the welfare gain. The gain is measured by per capita expenditure, associated with participation in the program, relative to the *status quo*. If farmer $i = \{1, 2, \ldots, N\}$ chooses to participate ($D_i = 1$), the relevant household outcome is Y_{1i}; Y_{0i} is the relevant outcome for non-participating ($D_i = 0$) households. Therefore, in regression format, $Y_i = Y_{0i} + D_i(Y_{1i} - Y_{0i}) + \eta_i = \alpha + \tau D_i + \eta_i$. Since participation is voluntary, the outcome is not likely to be independent of the treatment choice, such that treatment is not independent of the error term in the simple regression. Therefore, additional assumptions are needed.

3.1. Matching

Assuming that the distribution of welfare outcomes, Y_{1i} and Y_{0i}, are independent of treatment D_i, given a vector of covariates X_i, a propensity score matching estimator for the average effect of treatment on the treated can be derived. Intuitively, the goal of matching is to create a control group of non-PFM participants that is as similar as possible to the treatment group of PFM participants, although the groups differ in terms of their participation. Identification of the average effect, via matching, requires both strict ignorability of treatment, $(Y_{1i}\ Y_{0i}) \perp D_i | P(X_i)$, and propensity score overlap, $0 < P(X_i) < 1$ (Rosenbaum and Rubin, 1983, Dehejia and Wahaba, 1999, and Dehejia and Wahaba, 2002). The first assumes away the potential correlation between the treatment indicator and the error term. The second assumption results in a common support, in which similar individuals have a positive probability of being both participants and non-participants (Heckman et al., 1999). The matching analysis, below, considers nearest neighbor matches – we allow for between one and five matches – caliper matches and kernel matches. Both NN matches and caliper matches share the common feature of using only a few observations from the comparison group to construct treatment counterfactuals. Kernel matching, which uses a non-parametric weighting algorithm, provides an alternative, in which the weight is proportional to the propensity score distance between the treated and untreated. The advantage of kernel matching is greater efficiency, as more information is used; however, the disadvantage is that matching quality may be limited, due to use of observations that may be bad matches (Caliendo and Kopeinig, 2005).

3.2. Non-ignorable Treatment Assignment

If there are unobservable determinants of participation, meaning that treatment assignment is non-ignorable, matching estimators will be biased; therefore, control functions (selectivity models) or IV approaches are, instead, needed (Wooldridge, 2002, Heckman, and Navarro-Lozano, 2004, Heckman and Vytlacil, 2005 and Todd, 2008). We consider two different specifications in our analysis. The first specification we consider is 2SLS, in which instruments related to eligibility, along with household level variables, are used to control for participation. Given the prevalence of 2SLS in econometric textbooks, the model will not be described here. However, two concerns related to 2SLS are worth discussing: weak instruments and overidentification. A 2SLS model that contains weak instruments is not identified, and, therefore, the resulting treatment effects are not valid. Stock and Yogo (2005) tests for weak instruments were applied to consider this possibility. Furthermore, if the instruments affect participants in different ways, interpreting the resulting treatment effects estimates can be complicated. The IV treatment effect literature refers to this complication as treatment effect heterogeneity (Imbens and Angrist, 1994, Angrist et al., 1996, Frölich, 2007, Heckman and Vytlacil, 2005 and Todd, 2008). An overidentified model would be suggestive of treatment effect heterogeneity, and, therefore, we test our 2SLS models for overidentification. The second specification we consider is an extension of Heckman's (1979) selection model, due to Dubin and McFadden (1984) and Bourguignon et al. (2007), in which control functions are used to consider potential treatment effect heterogeneity. The program that we consider has potentially two levels of selection. In one level, households are deemed to be eligible, based on the presence of people from the Menja tribe. In another level, eligible households are allowed to decide whether or not they want to participate, which is likely to depend on their expectation of potential gains from participation.

In the case of sequential hurdles, such as these, a single-choice first stage, as is implied by 2SLS, is inefficient, since it does not control for correlations arising from unobservable components across the two choices (Khanna, 2001). To counter this limitation, a variety of polychotomous selection models has been proposed by Lee (1983), Dubin and McFadden (1984), Dahl (2002) and Bourguignon et al. (2007) to deal with this possibility. We follow Bourguignon et al.'s (2007) extension to Dubin and McFadden (1984), in which a multinomial logit of participation is estimated, followed by an OLS estimate of welfare that corrects for self-selection. The approach has been widely used in applied work related to, for example, climate change adaptation (Mansur et al., 2008) and technology adoption in agriculture (Caswell and Zilberman, 1985, and Wu and Babcock, 1998). Although the multinomial logit potentially suffers from violating independence of irrelevant alternatives (Khanna, 2001), it has been found to be robust in the face of various data generating processes, compared to its rivals (Bourguignon et al., 2007, Mansur et al., 2008).

3.3 Instruments

In this study a combination of instrumental variables were employed. One set of instruments relates to eligibility, due to the presence of the Menja tribe. This set includes a dummy indicator for the presence of individuals from the Menja tribe, and provides an indication of the intention to treat. In addition to this dummy indicator, the population density of the Menja tribe, and the total number of Menja households in the village, were also used as indicators of the intention to treat. As noted earlier, the Menja tribe was an important attribute of the forestry selection process, which further resulted in the provision of training with regard to the PFM. However, program eligibility was associated with deforestation, which could be associated with reduced household income and per capita consumption; therefore, the intention to treat IV implied by these eligibility instruments could be related to the outcome, which would bias the treatment effects estimates towards zero. For an upward bias to obtain, the presence of the Menja tribe would need to be associated with better welfare outcomes for eligible households than ineligible households.

Given this potential bias, household level instruments were also included in the analysis, such as the distance of the household from the PFM forest, the availability of alternative forests, and participation in other collective action programs. Households that have to travel farther to use a program forest should be less inclined to participate, which is also true for households that have access to an alternative forest. Similarly, households that have participated in other collective action programs could be more inclined to participate in another program, although it is possible that previous negative experiences would reduce the inclination to participate. The exclusion restriction related to these instruments may, however, be more contentious. For example, as was the case with the eligibility instruments, households farther away from program forests may be located near forests that are less degraded. Importantly, in a simple regression of per capita consumption against forest cover, access to alternative forests and previous participation in other collective action programs, not reported here, no correlation was uncovered, providing evidence in favor of these exclusion restrictions at the household level.

4. THE DATA

Data for the analysis was obtained from a household survey, designed for this study, undertaken in 10 Ethiopian villages in October of 2009. The villages are located in the Gimbo District, in southwestern Ethiopia. Sample frames for the survey were derived from the selected villages, via the lower level of local government, the kebele. The analysis was based on randomly selected households: 200 from PFM villages and 177 from non-PFM villages. Table 1 outlines the kebeles and the villages within the kebeles, including PFM participants and non-participants, and the number of survey respondents in each.

Respondents provided information on household characteristics, such as: age, education, gender, family size, household expenditure on various goods and services, household assets, household earnings from the sale of various goods and services, labor allocated to off-farm activities, distance to the nearest town and distance to the nearest road. Additional information related to potential determinants of PFM participation was also collected, including: the presence of members of the Menja tribe, total number of Menja households in the village, the Menja population density, distance from the PFM forest, availability of alternative forests and experience with other collective action arrangements. Finally, data related to the community, especially forest cover and population, was gathered.[9]

Descriptive statistics of that data are presented in Table 2, and these statistics are separated by participation status and eligibility. As expected, per capita expenditure is larger for the participating households, relative to non-participating households; similarly, it is higher for eligible households, relative to ineligible households. Given the way the program was developed, it is not surprising that eligible households are located in areas surrounded by the Menja tribe, although it is somewhat surprising that non-participating, but eligible, households are located in areas with heavier concentrations of the Menja tribe. It is also not surprising that eligible participating households were more likely to have had experience with other collective action programs, are located closer to program forests, are less likely to have access to alternative forests, are surrounded by poorer forest conditions and live in more

[9] In some cases, incomplete data was available. Where the number of incomplete cases was large, a dummy indicator of missing information was included, while the missing values were set to zero; this is the case for both population in the village and education in the household. Where the number of incomplete cases was small, the household was not included in the analysis, leading to approximately 370 observations, rather than the initial 377.

populated villages than either eligible non-participants and ineligible households. Taken together, these results suggest that our instrumentation strategy can be expected to perform adequately. Although eligible participating households appear to have smaller landholdings, and, therefore, less livestock, they also appear to be more active in off-farm activities than all other household types. Ineligible households, on the other hand, are generally located closer to both roads and town and are more educated than eligible participant and non-participant households.

For the sake of this study, per capita consumption expenditure, including goods produced at home, which were valued at village prices, was used to measure welfare, and not income, for the following reasons. First, by virtue of consumption smoothing, consumption expenditure fluctuates less in the short run compared to income. Second, consumption expenditure provides information over the consumption bundle that fits within the household's budget, although credit market access and household savings affect that budget (Skoufias and Katatyama, 2011); similarly, it is easily interpreted and widely used. As such, consumption is generally believed to provide better evidence of the standard of living than income. Third, an income survey may not capture informal, in-kind or seasonal income, and, thus, may be more susceptible to under-reporting. Unfortunately, the choice of per capita expenditure is not without problems. It might be preferred to measure it in adult equivalence, which takes into account differences between children and adults, in terms of their nutritional and other requirements. However, inaccuracies in adult equivalence would result in sizable measurement errors, limiting its usefulness.

5. RESULTS AND DISCUSSION

As noted earlier, if treatment assignment was completely random, it would be possible to simply compare the mean difference in per capita consumption. Since participation is voluntary, and, therefore, random treatment assignment may not obtain, we consider conditional mean differences based on matching, as well as instrumentation. Each is considered, in turn, below.

5.1. Matching

Before turning to the results, the underlying premises of matching – unconfoundedness and overlap – must be considered. Table 2, previously discussed, alludes to an initial balance test, suggesting wide differences between participating and non-participating households.

Therefore, in order to match and balance the data, program participation was estimated via logit regression. Propensity scores, the predicted probabilities of participation, were used as the matching basis. The logit results, presented in Appendix Table A.1, offer rather similar conclusions to those implied from comparing covariate mean differences, although the ability to simultaneously control for multiple covariates within the regression does yield some differences.

Since a wide range of matches is considered in the analysis, the match quality across these different algorithms deserves attention. The final choice of the matching algorithm is potentially guided by a broad set of criteria, primarily concerned with the quality of the match. Table A.2 outlines that quality.[10] One approach is to check if significant mean differences remain across the covariates, after matching. Another approach, suggested by Sianesi (2004), is to re-estimate the logit regression using the matched sample. After matching, there should be no systematic difference between covariates, and, thus, the pseudo-R^2 should be fairly low (Caliendo and Kopeinig, 2008). In the same vein, a likelihood ratio test of joint significance can be performed. The null hypothesis of joint insignificance should be rejected before matching, but not after matching. According to the results reported in Table A.2, four of the five NN matches resulted in balance for all of the covariates, as did one of the kernel matching algorithms. Furthermore, matched sample sizes were largest for the NN matches. Therefore, based on balancing, NN(2) through NN(5),[11] as well as kernel matching with a bandwidth of 0.0025, perform the best.

Although a subset of the proposed matching estimators perform better than the others, match-based treatment effects for all match algorithms were estimated. The treatment effects are available in Table 3. The first row reports simple mean differences. This naïve estimate suggests that there is a positive, but statistically insignificant welfare benefit. However, after controlling for program participation, via matching, the conclusion changes. For the best matches, NN(2)-NN(5) and 0.0025 bandwidth kernel matching, the program's average impact on program participants is estimated to range from ETB295.68 to ETB 548.53, and

[10] The important columns are the second and third columns. As 14 variables were included in the analysis, a test result of 14 in the second column suggests that the matching yields complete balance. The numbers in that column represent the number of insignificant mean differences, after matching. Furthermore, the pseudo-R^2 results contained in column 3 suggest that, with two exceptions (caliper = 0.01 and kernel bandwidth = 0.01), the re-estimated propensity score models have very limited explanatory power.
[11] NN(2), for example, refers to an algorithm that includes the two nearest matches.

each of the estimates are statistically significantly different from zero.[12] Given that average per capita expenditure for participating households is approximately ETB1732.09, the program impact accounts for between 17.8% and 31.7% of observed participant household per capita expenditure.

By the aforementioned standards, a number of matches perform rather well, but it should be noted that matching is based on an intrinsically non-testable assumption, conditional independence (Becker and Caliendo, 2007). However, if treatment assignment is non-ignorable, conditional independence is not appropriate, and match-based treatment effects are biased. The sensitivity of the estimates to uncontrolled bias could be either large or small (Rosenbaum, 2005). Although it is impossible to estimate the magnitude of the bias, it is possible to test the robustness of the matching estimates to potential unobserved variables. Rosenbaum's (2005) bounding approach is used in this analysis to examine the sensitivity of the match-based treatment effects estimates with respect to potential deviations from conditional independence. The results of that sensitivity analysis are presented in Table A.3. The first column of the table contains an odds ratio measure of the degree of departure from the outcome that is assumed to be free of unobserved bias, Γ.[13] The second column contains the upper bound p-value from Wilcoxon sign-rank tests examining the matched-based treatment effect for each measure of unobservable potential selection bias. As the estimated ATT values are positive, discussed below, the lower bound, which corresponds to the assumption that the true ATT has been underestimated, is less interesting (Becker and Caliendo, 2007), and is not reported in the table. From the table, we see that unobserved covariates would cause the odds ratio of treatment assignment to differ between the treatment

[12] The present value of the estimate multiplied by the size of relevant population yields the total benefit of the program interventions. Although comparing this quantity with the program cost would allow us to evaluate the cost-effectiveness of the program, we do not know the cost of the program for this region. However, the results could be used to create a cost-effectiveness measure, if the evaluator was willing to assume that the treatment effect was constant across the entire population. Furthermore, the results can be used to encourage participation amongst previously skeptical rural households.

[13] For ease of exposition, assume program participation is $P(x_i, u_i) = P(D_i = 1|x_i, u_i) = e^{\beta x + \gamma u_i}$. Therefore, the odds that two matched individuals, say m and n, receive the treatment may be written as $e^{\gamma(u_m - u_n)}$. Thus, two individuals with the same observable covariates may have differing program participation odds, due to differing unobserved effects, and the odds are influenced by the factor γ. If there is no difference in unobservable covariates or if these covariates don't affect participation, treatment assignment is random, conditional on the covariates. Thus, the Rosenbaum test assesses the required strength of γ or $u_m - u_n$ to nullify the matching assumption. Placing the condition within bounds, yields $e^{-\gamma} \leq e^{\gamma(u_m - u_n)} \leq e^{\gamma}$, implying that e^{γ} can be used to assess that strength. For example, if $\gamma = 0$, $e^{\gamma} = 1$, or $\Gamma = 1$, which implies that there is no problem. If, on the other hand, $\Gamma = 2$, one subject is twice as likely as another to receive the treatment, because of unobserved pretreatment differences As such, Γ measures the degree of departure from the random treatment assignment assumption that is inherent in matching (Kassie et al., 2011). If departure occurs at Γ values near 1, the matching estimate is highly sensitive to potential unobserved effects (Rosenbaum, 2005).

and control groups, once we reach a factor of about 1.7. Therefore, we conclude that there is strong evidence that the matching method estimates are somewhat sensitive to selectivity bias. However, as Becker and Caliendo (2007) note, this sensitivity result is a worst-case scenario. It does not test for unobserved factors; rather, it indicates that the program effect confidence interval would include zero, if the unobserved covariates cause the program participation odds to differ by a factor of 1.7.

5.2. Instrumentation via Two-Stage Least Squares

Given the fact that participation could be determined by both observed and unobserved factors, matching estimates are likely to be biased, and, as outlined in the discussion of the program in Section 2, participation is a potentially complex scenario including both village eligibility and household decisions. Therefore, the analysis was extended to allow for potential unobserved factors that could affect participation. In the initial analysis, treatment effects were examined based on an intention to treat strategy, whereby variables associated with the menja population – whether or not menja were present in the village, the density of the menja population in the village, and the total number of menja households in the village – were used as instruments for participation. The primary results from that analysis are reported in Table 4. In a follow-up analysis, treatment effects were examined based on household level decisions, such that the distance the PFM forest, the availability of an alternative forest, and household experience with previous collective action programs were used to instrument for participation. Results from that analysis are available in Table 5. The preceding two analyses were extended to consider various combinations of instruments both for eligibility and for household decisions. However, only results considering all eligibility instruments combined with each of the household decision instruments are reported in Table 6. The full set of second-stage regression results for the model associated with Table 6 is presented in Table 7.[14] The analysis was performed using software developed by Baum et al. (2010).

Tables 4-6 contain only a subset of the results from the various 2SLS specifications. The top panel of each table contains the estimated treatment effects and the R^2 from the model, while the second panel contains the estimates of the different excluded variables on participation, along with the R^2 from the first stage. In the third and final panel of each table, the performance of the 2SLS model is outlined, and, thus, includes a variety of performance

[14] Results from different combinations are available from the authors, but the results are very similar to those discussed, and, therefore, are not provided here.

statistics, including: Kleibergen and Paap's (2006) LM statistic (KP LM) of model fit, Cragg and Donald's (1993) Wald statistic (CD Wald) of model identification, Stock and Yogo's (2005) 10% critical value (SY 10%) of weak instrumentation, which is applied to the CD Wald value, the Anderson and Rubin (1950) Wald statistic for the hypothesis that the treatment effect is both insignificant and exogenously identified by the instruments, and Hansen's J-statistic of overidentification, when more than one instrument is included. In all models, standard errors are clustered at the village level to control for potential correlation within the villages.

According to the t-statistics, treatment effects are estimated to be positive and significant in all but three of the reported analyses, and the effect ranges between ETB370.00 and ETB 468.80. Of the remaining three insignificant estimates, two were negative (ETB -38.61 and ETB -96.05), and one was well above the range (ETB 619.20). In the case of the two negative and insignificant estimates, the Anderson-Rubin test rejected either the significance of the estimate or the exogeneity of the instruments, while the Stock-Yogo test cannot reject that the instruments used in the analysis are too weak, at the 10% level, to be considered valid. However, in the remaining analyses, instruments were valid, given the Stock-Yogo 10% critical value. Finally, in all cases, the models were not overidentified, regardless of the number of instruments used.

Our preferred estimates are contained in Tables 6 and 7, and are based on the inclusion of all eligibility instruments and household decision instruments. Table 6 contains the subset of results for both the first and second stage, while Table 7 reports the full set of estimates from the second stage. The estimated treatment effects in our preferred models are significant and range between ETB 377.60 and ETB 400.60, with standard errors in the region of ETB 170.00, such that a 95% confidence interval of the treatment effect falls between ETB 40.00 and ETB 750.00. Given that the mean per capita consumption of participants is ETB 1732.09, as noted above, the program has increased per capita consumption by either as little as 2.3% or by as much as 43.3%, with an average increase of approximately 22.5%. Not only are the treatment effects statistically significant, they are economically significant, as well.

In addition to the participation effects previously discussed, a number of other variables are found to be important determinants of per capita consumption, and the discussion of those variables is based on Table 7.[15] As would be expected, simply from the definition of per capita consumption, larger households spend approximately ETB 192 less per household member. Households with corrugated roofs (significant), greater landholdings (insignificant), family members participating in off-farm earning opportunities (insignificant), and families with more livestock (significant) consume significantly more per capita. Adding the estimates across these categories, while assuming independence (for purposes of standard error calculation), suggests a significant increase in per capita consumption of ETB 430 (s.e. = 172.08). Households with poorer access to markets, located either farther from roads or from towns, consume significantly less: ETB 5.99 (2.07). More education in the household is associated with significantly higher per capita expenditure: ETB 37.05 (18.92), while per capita consumption is significantly lower in both larger villages (ETB 0.12 per villager) and in villages with missing population figures (ETB 518-525 per village). Current forest cover was not found to be a significant determinant of per capita expenditure, although better forest cover was associated with an economically significant increase in per capita consumption of ETB 174-199.

5.3. Multilevel Selection

Although the results from 2SLS are rather robust to instrument choice, it is possible that the eligibility effect and the household decision effect work in opposite directions. Therefore, the analysis was extended to consider the possibility that once eligibility was determined, selection at the household level could have affected the results. In order to consider the possibility, the data was split into three groups: ineligible households, eligible participating households and eligible non-participating households. The probability of a household falling into any of the three groups was estimated via multinomial logit, and the effect of selection on household welfare was estimated via least squares regression. The model structure was based on an alternative to Dubin and McFadden (1984), developed by Bourguignon et al. (2007), which allows the selection effects to be free, meaning that they are not restricted to sum to zero. Bourguignon et al.'s (2007) Monte Carlo simulations suggest that not imposing this restriction does not result in a loss of efficiency, even if the restriction is not valid.

The results from this specification are reported in Table 8. In the first two columns, the multinomial logit parameter estimates are presented; as there are only three groups, the

[15] Other second-stage regression results are also available from the authors. However, those results are very similar to the results reported in Table 7, and, therefore, they are not reported here.

estimates are relative to the left-out category, eligible and participating households. For instruments, only household level determinants were used, such as distance to the PFM forest, whether or not an alternative forest was available, and whether or not the household had previous experience with other collective action programs.[16] The final three columns of the table present estimates for the mean level of per capita consumption for each category of participant, conditioned on household level factors (the same as those used in the 2SLS models discussed above) and the different selection effects. The relevant rows within each of the last three columns relate to these selection effect estimates. Despite the fact that the three groups were empirically separable, given the instruments used (see the first few rows of the first two columns), all of the selection effects were imprecisely estimated. In other words, we cannot reject the hypothesis that, once eligibility is conditioned, household participation decisions are not correlated with the error term in the regression. Although it may seem surprising that there are no selection effects, other than eligibility selection effects, these results are, for the most part, in agreement with those reported in Tables 4-6, which found rather similar treatment effects estimates, regardless of which set of instruments were included. Furthermore, as noted above, regardless of which set of instruments were included, overidentification was rejected in each of the 2SLS models.

5.4 Multilevel Matching

6. CONCLUSION

Previous studies that have evaluated the welfare impacts of common property forestry programs, have found a wide variety of results that depend upon the study context and the employed methodology. Motivated by these uncertainties, the present study set out to evaluate the welfare impact of a common property forestry program that resulted in the decentralization of forestry management and was augmented by market linkage interventions. The analysis was based on data collected in selected villages of the Gimbo district in southwestern Ethiopia. The potential outcome framework underpinned the analysis, through which, the causal link between program intervention and household welfare could be empirically investigated. Compared to related program evaluations, such as that by Jumbe

[16] Eligibility is strictly defined by variables associated with the Menja tribe, such that these could not be included in the first-stage MNL regression. Specifically, inclusion of any of the Menja variables led to a Pseudo-R^2 of unity, and second-stage regressions showing strong signs of multicollinearity.

and Angelsen (2006) and Cooper (2008), we considered matching, instrumental variables and selection models to examine the robustness of the estimated treatment effects to various identification assumptions. The analysis revealed that this PFM forestry intervention has raised the average welfare of participating households in the study villages. Identification based on observed controls, via matching, yielded an average program effect of approximately ETB336.73. However, this matching estimate was generally lower than the program effect estimated via instrumental variables methods controlling for both observed and unobserved factors. Our 2SLS estimates ranged from ETB 370.60 to ETB 468.80, although the matching estimate falls within the 95% confidence interval of all of the IV model treatment effects estimates.

The results from the analysis imply that the decentralization of natural forest management, when combined with market access support for NTFPs, has substantially raised participant household welfare, accounting for approximately one-quarter of total welfare. In terms of policy, although, due to the program design, we are not able to directly attribute the gains to either the change in forestry management arrangements or market access, we are led to believe, on the basis of this study's findings, as well as the restrictions imposed on forestry product harvests in the PFM forests, that the welfare gain is most likely to be due to the market linkages associated with NTFPs harvested from PFM forests. The results from this analysis support Dasgupta (2006) and are consistent with the conjuncture proposed by Wunder (2001), that NTFP marketing offers sizeable welfare gains.

Although the benefits of the program are estimated to be large, and positive, this research has not considered program equity, which is likely to be an important determinant of the ability of these programs to further enhance rural development. If village elites are able to capture the majority of the benefits of the program, the resulting inequality could mitigate the perceived effectiveness of the program. Future research should consider the possibility that, even though the program has increased welfare, it may not have lifted all boats, equally, as has been found by Jumbe and Angelson (2006) and Cooper (2008).

REFERENCES

Agrawal, A. and Gibon, C. (1999). Enchantment and disenchantment: The role of community in natural resource conservation. *World Development, 27*: 629-49.

Agrawal, A., and Ostrom., E. (2001). Collective action, property right and decentralization in resource use in India and Nepal. *Politics and Society,* 4: 485-514.

Angelsen, A., and Wunder, S. (2003). Exploring forestry-poverty link: key concepts. Issues and research implications, CIFOR Occasional Paper, vol.4, CIFOR, Bogor.

Angrist, J., Imben, G., and Rubin. D. (1996). Identification of causal effects using instrumental variables. *Journal of American Statistical Association*, 91: 444-472.

Arnold, M., (1998). Forestry and sustainable livelihoods. In: Carney, D. (Ed.) Sustainable Rural Livelihoods: What Contributions Can We Make? Department for International Development, London, pp. 155–166.

Arnold, M., and Pérez. M. (2001). Can non-timber forest products match tropical forest conservation and development objectives?, *Ecological Economics*, 39: 437–447

Baum, C.F., Schaffer, M.E., Stillman, S. (2010). ivreg2: Stata module for extended instrumental variables/2SLS, GMM and AC/HAC, LIML and k-class regression. http://ideas.repec.org/c/boc/bocode/s425401.html.

Becker, S.O., and Caliendo, M. (2007). Sensitivity analysis for average treatment effects. *Stata Journal*, 7: 71–83.

Bekele, M., and Bekele, T. (2005). Participatory Forest Management in Chilimo and Bonga, Ethiopia. An Evaluation Report, Farm Africa, Addis Ababa.

Blomley, T., Pfliegner, K., Isango, J., Zahabu, E., Ahrends, A., and Burgess, N. (2008). Seeing the wood for the trees: An assessment of the impact of participatory forest management on forest conditions in Tanzania. *Fauna and Flora International* 42: 380-391.

Bluffstone, R. (2008). Does better common property forest management promote behavioral change? On farm tree planting in the Bolivian Andes. *Environment and Development Economics,* 13:137-170.

Bourguignon, F., Fournier, M., and Gurgand, M. (2007). Selection bias corrections based on the multinomial logit model: Monte Carlo comparisons. *Journal of Economic Surveys*, 21: 174-205.

Caliendo, M., and Kopeinig, S. (2005). Some practical guideline for the implementation of propensity score matching. *IZA Discussion Paper No. 1588*.

Caswell, M.E, and Zilberman, D. (1985). The choices of irrigation technologies in California. *American Journal of Agricultural Economics*, 67: 224-34.

Cobb-Clark, D.A., and Crossley, T. (2003). Econometrics for evaluation: An introduction to recent developments. *The Economic Record,* 79: 491-511.

Cooper, C. (2007). Distributional consideration of forest co-management in heterogeneous community: Theory and simulation. *Department of Economics Working Paper*, University of Southern California.

Cooper, C. (2008). Welfare effects of community forest management: Evidences from the hills of Nepal. *Department of Economics Working Paper*, University of Southern California.

Cragg, J.G. and Donald, S.G. (1993). Testing identfiability and specification in instrumental variables models. *Econometric Theory*, 9: 222-240.

Dahl, G.B. (2002). Mobility and the Returns to Education: Testing a Roy Model with Multiple Markets, *Econometrica* 70: 2367-2420.

Dasgupta, P. (2006). Common property resources as development drivers: A study of fruit cooperative in Himachal Pradesh, India. *SANDEE Working Paper No. 15-06*.

Dehejia, R. H., and Wahaba, S. (1999). Causal effects in non-experimental studies: Reevaluating the evaluation of training programs. *Journal of the American Statistical Association*, 94: 1053-1062

Dehejia, R. H., and Wahaba, S. (2002). Propensity score matching methods for non-experimental causal studies. *Review of Economics and Statistics,* 84: 151–61.

Delacote, P. (2007). Agricultural expansion, forest products as safety net and deforestation. *Environment and Development Economics*, 12: 235-249.

Dubin, J. A., McFadden, D. L. (1984). An Econometric Analysis of Residential Electric Appliance Holdings and Consumption, *Econometrica* 52: 345-362.

Edmonds, E. (2002). Government-initiated community resources management and local resource extraction from Nepal's forests. *Journal of Development Economics,* 68: 89-115.

Evans, M.I., (1993). Conservation by commercialization. In: Hladik, C.M., Hladik, A., Linares, O.F., Pagezy, H., Semple, A., Hadley, M. (Eds.), Tropical Forests, People and Food: Biocultural Interactions and Applications to Development. MAB Series, vol. 13'. UNESCO, Paris and Parthenon Publishing Group, Carnforth, UK, pp. 815 822.

Gobeze T., Bekele, M., Lemenih. M., and Kassa. H.(2009). Participatory forest management impacts on livelihood and forest status: The case of Bonga forest in Ethiopia. *International Forestry Review,* 11: 346-358.

Frölich, M. (2007). Nonparametric IV estimation of local average treatment effects with covariates. *Journal of Econometrics,* 139: 35-75.

Gauld, R. (2002). Maintaining centralized control in community based forestry: policy construction in the Philippines. *Development and Change,* 31: 229-254.

Heckman, J.J. (1979). Sample selection bias as a specification error. *Econometrica*, 47: 153-161.

Heckman, J.J., Ichimura, H., Smith, J., and Todd, P. (1998). Characterizing selection bias as using experimental data. *Econometrica,* 66: 11017-1098.

Heckman, J.J., LaLonde, R., Smith, J. (1999). The economics and econometrics of active labor market programs. *The Hand Book of Labor Economics,* North Holland, New York, pp: 1865-2097.

Heckman, J.J., and Navarro, S. (2004). Using matching, instrumental variables, and control functions to estimate economic choice models. *Review of Economics and Statistics*, 86: 30–57.

Heckman, J.J., and Vytlacil, E.J. (2005). Structural equations, treatment effects and econometric policy evaluation. *Econometrica,* 73: 669-738.

Imbens, G., and Angrist, J. (1994). Identification and estimation of local average treatment effects. *Econometrica,* 62: 467-475.

Jirane, T., Tadesse, T., and Temesgen, Z. (2008). PFM in Oromia and SNNP regions of Ethiopia: A review of experiences, constraints and implications for forest policy, in J. Bane, S. Nune, A. Mekonnen and R. Bluffstone (eds.), *Policies to increase forest cover in Ethiopia*, Proceedings of Policy Workshop, Environment for Development Environmental Economics Policy Forum for Ethiopia, Ethiopian Development Research Institute, Global Hotel, Addis Ababa.

Jumbe, C.B.L., and Angelsen, A. (2006). Do the poor benefit from devolution policies? Evidence from Malawi co-management programs. *Land Economics,* 84: 562–581.

Kassa, H., Campbell, B., Sandwall, M., Kebede, M., Tesfaye, Y., Dessie, G., Seifu, A., Tadesse, M., Garedew, E., and Sandwall, K. (2009). Building a future scenario and uncovering persisting challenges in Chilimo forest, Central Ethiopia. *Journal of Environmental Management,* 90: 1004-1013.

Khanna M. (2001). Sequential adoption of site-specific technologies and its implications for nitrogen productivity: A double selectivity model. *American Journal of Agricultural Economics*, 83: 35-51.

Kleibergen, F. and Paap, R. (2006). Generalized reduced rank tests using the singular value decomposition. *Journal of Econometrics*, 133: 97-126.

Klooster, D., and Masera, O. (2000). Community forest management in Mexico: Carbon mitigation and biodiversity conservation through rural development. *Global Environmental Change,* 10: 259-272.

Lee, L. F. (1983). Generalized Econometric Models with Selectivity, *Econometrica,* 51: 507-512.

Lemenih, M., and Bekele, M. (2008). Participatory forest management, best practices, lessons and challenges encountered: The Ethiopian and Tanzanian Experiences. An Evaluation Report, Farm Africa, Addis Ababa.

Mansur E.T., Mendelsohn R., Morrison W. (2008). Climate change adaptation: A study of fuel choice and consumption in the US energy sector, *Journal of Environmental Economics and Management,* 55: 175–193.

Mekonnen, A. (1999). Rural household biomass fuel production and consumption in Ethiopia: A case study. *Journal of Forest Economics* 5:69-97.

Mekonnen, A., and Bluffstone, R. (2008). Policies to increase forest cover in Ethiopia: lessons from economics and international experience. in J. Bane, S. Nune, A. Mekonnen and R. Bluffstone (eds.), *Policies to increase forest cover in Ethiopia*, Proceedings of Policy Workshop, and Environment for Development-Environmental Economics Policy Forum Ethiopia, Ethiopian Development Research Institute, Global Hotel, Addis Ababa.

Murty, M.N. (1994). Management of common property resources: Limits to voluntary collective action, *Environment and Resources Economics,* 4: 581-594.

Nagendra, H. (2002). Tenure and forest conditions: Community forestry in the Nepal Terai, *Environmental Conservation,* 29: 530-539.

Neumann, M. (2008). Participatory forest management in Oromia region of Ethiopia: a review of experience, constraints and implications for forest policy. in J. Bane, S. Nune, A. Mekonnen and R. Bluffstone (eds.), *Policies to increase forest cover in Ethiopia*, Proceeding of Policy Workshop, and Environment for Development-Environmental Economics Policy Forum Ethiopia, Ethiopian Development Research Institute, Global Hotel, Addis Ababa.

Neumann, R.P. and E. Hirsch (2000). Commercialization of non-timber forest products: review and analysis of research, Center for International Forestry Research (CIFOR), Bogor, Indonesia.

Nune, S. (2008). Ethiopian Government efforts to increase forest cover: a policy oriented discussion paper, in J. Bane, S. Nune, A. Mekonnen and R. Bluffstone (eds.), Policies to increase forest cover in Ethiopia, Proceeding of Policy Workshop, Environment for Development—Environmental Economics Policy Forum for Ethiopia, Ethiopian Development Research Institute, Global Hotel, Addis Ababa.

Robinson, E.J.Z., and Lokina, R.B. (2012). Efficiency, enforcement and revenue tradeoffs in participatory forest management: an example from Tanzania, *Environment and Development Economics,* 17: 1–20

Rosenbaum, P.R., and Rubin, D.B. (1983). The central role of propensity score in observational studies for causal effects. *Biometrika*, 70: 41-75.

Rosenbaum, R.P. (2005). Sensitivity analysis in observational studies. *Encyclopedia of Statistics in Behavioral Science,* 4: 1809–1814.

Roy, A.D. (1951). Some Thoughts on the Distribution of Income. *Oxford Economic Papers* 3: 135–46.

Rubin, D. (1973). Matching to remove bias in observational studies, *Biometrics*, 29: 159-183.

Shillington, L.J. (2002). Non-timber Forest Products, Gender, and Households in Nicaragua: A Commodity Chain Analysis, MSc Thesis, Virginia Polytechnic Institute and State University.

Shumeta, Z., Urgessa, K., and Kebebew, Z. (2012). Analysis of market chains of forest coffee in southwest Ethiopia, *Academic Journal of Plant Sciences*, 5: 28-39.

Sianesi, B. (2004). An evaluation of the active labor market programs in Sweden. *Review of Economics and Statistics,* 86: 133-15.

Skoufias, E., and Katayama, R.S. (2011). Sources of welfare disparities between and within regions of Brazil: Evidences from the 2002-2003 household budget survey (POF), *Journal of Economic Geography*, 11: 897-918.

Stock, J. H., and Yogo, M. 2005. Testing for Weak Instruments in Linear IV Regression. In D.W.K. Andrews and J.H. Stock, eds. Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg. Cambridge: Cambridge University Press, 2005, pp. 80-108. Working paper version: NBER Technical Working Paper 284. http://www.nber.org/papers/T0284.

Sunderlin, W., Angelsen, A., Belcher, B., Burger, P., Nasi, R., Santoso, L., and Wunder, S., (2005). Livelihood, Forests and conservation in developing countries. *World Development,* 9: 1383-1402.

Todd, P. (2008). Evaluating social programs with endogenous program placement and selection of the treated, Handbook of Development Economics, Volume 4:pp *848-873*

Wooldridge, J. M. (2002). Econometric Analysis of Cross Section and Panel Data, London: The MIT Press.

Wu, J., and Babcock, B.A. (1998). The Choice of Tillage, Rotation and Soil Testing Practices: Economic and Environmental Implications, *American Journal of Agricultural Economics,* 80: 494-511.

World Bank. (2005). The Little Green Data Book. Washington, DC: World Bank.

World Bank. (2008). Poverty and the environment: Understanding the linkage at the household level. Washington DC, World Bank.

WDR. "World Development Report" (2008): Agriculture for Development.

Wunder, S. (2001). Poverty alleviation and tropical forests – what scope for synergies? *World Development,* 11: 1817-1833.

Yadev, N.P., Dev, O.P., Springate-Baginski O. and Sousan, J. (2003). Forest management and utilization under community forestry. *Journal of Forests and Livelihood,* 1: 37-50.

Table 1. Sample Villages and Respective Sample Sizes

List of Kebeles	Number of villages	Name of villages	
		PFM villages	Non-PFM villages
Yebito (88)	2	Agama (58)	Mula - Hindata (30)
Bita Chega(49)	1	Dara (49)	
Michiti (80)	3	Beka (32), Matapha (24)	Chira - Botera (24)
Woka Araba (50)	1		Woka-Araba (50)
Keja Araba (47)	1		Keja-Araba (47)
Maligawa (63)	2	Sheka (37)	Sheko (26)
Total	10	200	177

Table 2. Descriptive Statistics for Baseline Covariates and Household Welfare

Variable Definition	Non-participant	Eligible Participant	Eligible Non-participant
Per capita expenditure in Ethiopian Birr	1,624	1,832	1,786
	(112.8)	(159.0)	(24.3)
Dummy: Menja households in community	0	1	1
Menja households, % of total	0	0.238	0.529
		(0.17)	(0.19
Total number of menja households	0	34.15	67.99
		(14.0)	(25.4)
Dummy: Household ever participated in other collective action	0.027	0.16	0.093
	(0.02)	(0.03)	(0.00)
Distance to programme forest in minutes	90.53	23.63	36.16
	(35.4)	(3.91)	(1.56)
Dummy: Access to alternative forest	0.505	0.222	0.453
	(0.08)	(0.07)	(0.05)
Dummy: Household member participated in off-farm employment	0.108	0.153	0.04
	(0.06)	(0.05)	(0.01)
Household landholding in hectares	2.499	2.168	2.292
	(0.22)	(0.15)	(0.34)
Age of household head	43.41	43.33	43.08
	(3.53)	(1.62)	(1.14)
Total number of household members	5.838	5.84	5.733
	(0.32)	(0.26)	(0.16)
Dummy: House has corrugated roofing	0.306	0.25	0.173
	(0.08)	(0.07)	(0.03)
Distance to town in minutes	62.57	76.56	87.88
	(1.87)	(9.47)	(2.42)
Distance to nearest road in minutes	22.62	24.5	47.91
	(9.80)	(9.32)	(24.25)
Household livestock ownership converted to TLU	4.524	4.197	4.472
	(0.21)	(0.18)	(0.40)
Maximum education of any household member	6.622	5.514	5.507
	(0.96)	(0.36)	(0.53)
Maximum education mission	0.045	0.111	0.147
	(0.02)	(0.03)	(0.07)
Dummy: Forest cover is good	0.577	0.16	0.987
	(0.37)	(0.17)	(0.02)
Village population	1,932	2,149	411.3
	(903)	(467)	(250)
Village population not available	0.207	0.153	0
	(0.25)	(0.17)	
Observations	111	144	75

Robust (clustered at the village level) standard errors in parentheses.

Table 3. Matching-Based Treatment Effect on Treated Estimates of Welfare Effects

Estimator	ATT	Standard Deviation	t-Statistics
Simple mean difference	45.397	88.85	0.51
Nearest neighbor(1)	359.35	131.56	2.73 ***
Nearest neighbor (2)	295.68	111.87	2.64***
Nearest neighbor (3)	336.73	101.53	3.32***
Nearest neighbor(4)	327.62	105.30	3.11***
Nearest neighbor (5)	319.95	101.91	3.14
Radius matching(r=0.01)	103.17	1070	0.09
Radius matching(r=0025)	548.53	148.61	3.69**
Radius matching (r=0.005)	548.53	150.92	3.63**
Kernel matching(bwdth=0.01)	103.17	1150	0.09
Kernel matching(bwdth=0.0025)	548.53	152.84	3.58***
Kernel matching(bwdth=0.005)	548.53	154.91	3.54***

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 4. 2SLS Estimates of Model I: Menja Exclusions

VARIABLES	2SLS:Ia	2SLS:Ib	2SLS:Ic	2SLS:Id
Second-Stage Regression				
Treatment Effect	419.1**	619.2	440.0**	382.0**
	(165.0)	(387.4)	(223.1)	(172.0)
R-squared	0.241	0.217	0.239	0.244
First-Stage Regression				
Menja Indicator	0.635***			0.660***
	(0.111)			(0.094)
Menja Density		0.456		-0.658**
		(0.287)		(0.237)
Menja Households			0.0090**	0.0054*
			(0.029)	(0.034)
R-squared	0.795	0.550	0.677	0.838
Observations	370	370	369	369
2SLS Performance				
KP LM	127.78***	66.60***	100.29***	127.99***
CD Wald	512.28	40.5	195.08	247.11
SY 10%	16.38	16.38	16.38	22.30
AR Wald	8.93***	2.71*	5.10**	3.03**
Hansen J				0.485

Robust (clustered by village) standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 5. 2SLS Estimates of Model II: Household Exclusions

	Second-Stage Regression			
VARIABLES	2SLS:IIa	2SLS:IIb	2SLS:IIc	2SLS:IId
Treatment Effect	-38.61	468.8**	-96.05	370.0**
	(489.7)	(198.0)	(500.3)	(182.1)
R-squared	0.243	0.236	0.240	0.249
	First-Stage Regression			
Other Collective Actions	0.229**			0.222**
	(0.101)			(0.079)
Distance Program Forest		-0.0035***		-0.0034***
		(0.0004)		(0.0003)
Distance Other Forest			-0.165**	-0.0698
			(0.055)	(0.045)
R-squared	0.517	0.643	0.525	0.670
Observations	370	370	369	369
	2SLS Performance			
KP LM	15.92***	52.21***	16.99***	65.70***
CD Wald	13.28	143.48	17.34	59.89
SY 10%	16.38	16.38	16.38	22.3
AR Wald	0.00	5.86**	0.04	1.94
Hansen J				1.348

Robust (clustered by village) standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 6. 2SLS Estimates of Model III: All Menja and Household Exclusions

Second-Stage Regression				
VARIABLES	2SLS:IIIa	2SLS:IIIb	2SLS:IIIc	2SLS:IIId
Treatment Effect	377.6**	390.1**	400.6**	400.7**
	(170.9)	(169.2)	(178.9)	(173.9)
R-squared	0.244	0.243	0.247	0.247
First-Stage Regression				
Menja Indicator	0.657***	0.608***	0.650***	0.596***
	(0.0936)	(0.108)	(0.0926)	(0.106)
Menja Density	-0.651**	-0.606**	-0.631**	-0.575**
	(0.235)	(0.234)	(0.236)	(0.229)
Menja Households	0.0052*	0.0049*	0.0051*	0.0045*
	(0.002)	(0.002)	(0.002)	(0.002)
Other Collective Action	0.0448**			0.0612***
	(0.019)			(0.018)
Distance Program Forest		-0.0008		-0.0008
		(0.0005)		(0.0005)
Distance Other Forest			-0.0531**	-0.0407
			(0.023)	(0.027)
R-squared	0.839	0.844	0.840	0.846
Observations	370	370	369	369
2SLS Performance				
KP LM	128.70***	128.62***	128.40***	129.20***
CD Wald	185.93	194.04	185.82	129.91
SY 10%	24.58	24.58	24.58	29.18
AR Wald	2.45**	2.87**	2.67**	2.06*
Hansen J	1.291	1.072	1.380	2.019

Robust (clustered by village) standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 7. Second Stage Estimates from 2SLS

VARIABLES	Model IIIa	Model IIIb	Model IIIc	Model IIId
Treatment Effect	377.6**	390.1**	400.6**	400.7**
	(170.9)	(169.2)	(178.9)	(173.9)
Off-farm Labour	186.3	186.0	191.8	191.8
	(144.0)	(144.0)	(143.1)	(143.0)
Land Holdings	31.42	31.59	33.67	33.67
	(35.07)	(35.03)	(35.17)	(35.15)
Age HH Head	-7.088**	-7.104**	-7.171**	-7.171**
	(2.97)	(2.98)	(3.02)	(3.01)
HH size	-191.8***	-192.1***	-191.8***	-191.8***
	(15.35)	(15.47)	(15.60)	(15.59)
Corrugated Roofing	163.8*	163.9*	169.2**	169.2**
	(86.36)	(86.19)	(85.15)	(85.13)
Distance to Town Minuts	-2.423	-2.437	-2.471	-2.471
	(1.554)	(1.551)	(1.531)	(1.527)
Distance to Road Minutes	-3.571***	-3.584***	-3.714***	-3.714**
	(1.373)	(1.388)	(1.44)	(1.45)
Livestock TLU	49.75***	49.79***	49.00***	49.00***
	(15.95)	(15.98)	(16.19)	(16.19)
Education Max	37.05*	37.15*	35.01**	35.01**
	(18.92)	(18.96)	(17.64)	(17.67)
Education Missing	-39.34	-40.26	-46.34	-46.35
	(171.6)	(171.0)	(168.4)	(168.3)
Forest Cover	174.0	180.6	199.0	199.1
	(164.7)	(163.4)	(171.5)	(168.1)
Village Population	-0.123***	-0.125***	-0.125***	-0.125***
	(0.0382)	(0.0391)	(0.039)	(0.039)
Population Missing	-518.3***	-525.6***	-523.2***	-523.3***
	(184.4)	(185.2)	(185.3)	(183.9)
Constant	2,874***	2,872***	2,871***	2,871***
	(221.6)	(221.6)	(220.3)	(220.5)
Observations	370	370	369	369
R-squared	0.244	0.243	0.247	0.247

Robust (clustered by village) standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 8. Dubin-McFaddin Selection Model

VARIABLES	First-Stage MNL Ineligible	First-Stage MNL Eligible Nonparticipant	Second-Stage OLS Ineligible	Second-Stage OLS Eligible Participant	Second-Stage OLS Eligible Nonparticipant
Dist. Program	0.0434***	0.0201***			
Forest	(0.007)	(0.006)			
Dist Other	-0.0950	-0.790**			
Forest	(0.366)	(0.351)			
Collective	-4.959***	-2.891***			
Action	(1.077)	(1.055)			
Off-farm	0.189	-1.654**	-23.20	529.30**	-884.90
Labour	(0.812)	(0.691)	(306.5)	(241.1)	(606.2)
Land	-0.0061	-0.3100**	75.09	132.30	24.09
Holdings	(0.033)	(0.140)	(54.51)	(81.70)	(81.21)
Age HH	0.0039	0.0123	-16.75***	-7.96	-11.20
Head	(0.013)	(0.014)	(5.584)	(6.750)	(9.555)
HH Size	-0.0098	-0.1100	-156.2***	-218.0***	-176.9***
	(0.082)	(0.101)	(35.14)	(44.55)	(58.86)
Corr Roofing	-0.818***	-0.129	-3.431	350.1	7.294
	(0.313)	(0.505)	(162.5)	(218.2)	(330.7)
Dist Town	-0.0017	0.0221***	-1.711	-2.268	-2.427
	(0.005)	(0.005)	(2.962)	(1.760)	(2.907)
Dist Road	-0.0332	-0.0104	-4.285	-1.671	-1.399
	(0.023)	(0.018)	(4.198)	(3.956)	(3.522)
Livestock TLU	0.0030	0.0076	70.18**	-5.720	95.41**
	(0.062)	(0.077)	(33.84)	(28.62)	(40.65)
Educ Max	0.0631	-0.141**	55.70*	16.46	16.31
	(0.078)	(0.056)	(30.35)	(29.90)	(51.03)
Educ Missing	-0.838**	-1.265	-164.6	-176.7	21.03
	(0.413)	(0.776)	(256.1)	(304.6)	(492.2)
Forest Cover	4.118**	7.299***	-29.47	-396.8*	-670.1
	(1.845)	(1.564)	(243.0)	(240.8)	(1,026)
Vill Population	-0.0007	-0.0024***	0.0199	0.0827	0.0713
	(0.0004)	(0.0007)	(0.067)	(0.103)	(0.566)
Selection 1:			-195.0	121.7	-2,510
Ineligible			(412.8)	(1,202)	(3,553)
Selection 2:			-360.1	134.1	-1,846
Elig Part			(1,192)	(332.6)	(3,712)
Selection 3:			-977.5	339.5	-34.07
Inelig Part			(974.0)	(1,342)	(635.8)
Constant	-1.412	-1.980	2,215***	3,098***	1,316
	(2.418)	(2.178)	(390.9)	(695.2)	(4,106)

Standard errors in parentheses: MNL standard errors clustered by village, selection equation standard errors from 300 bootstrap replications; *N=330*. *** p<0.01, ** p<0.05, * p<0.1

A. PROPENSITY SCORE MATCHING REGRESSION AND SENSITITIVY

Table A.1. Propensity Score Estimates of Program Participation

VARIABLES	coefficient	Marginal effect
Household head's age	-0.008	-0.002
	(0.011)	(0.002)
Household head's gender	-0.336	-0.083
	(0.553)	(0.137)
Household head's education	0.022	0.005
	(0.052)	(0.012)
Female labour force	0.848***	0. 208***
	(0.307)	(0. 075)
Male labour force	-0.230	-0.056
	(0.258)	(0. 063)
Land holding size in ha	0.010	0.002
	(0.085)	(0.021)
Off-farm employment	0.842*	0.207*
	(0.490)	(0. 115)
Distance to agro extension office	-0.004*	-0.001*
	(0.002)	(0.001)
Woodlot ownership	-0.511*	-0.125*
	(0.282)	(0.068)
Livestock holding size in TLU	0.122***	0.030**
	(0.049)	(0. 012)
Distance from PFM forest	-0.028***	-0.006***
	(0.005)	(0.001)
Experience of other collective action	1.400***	0.329***
	(0.509)	(0 .103)
Distance from nearest town	-0.005*	-0.001*
	(0.003)	(0 .001)
Distance from nearest road	-0.008**	-0.002**
	(0.004)	(0.001)
Constant	0.281	
	(0.761)	
N	337	337

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A.2. Matching Estimator Performance

Matching estimator	Balancing test*	pseoudo-R2	matched sample size
Nearest-neighborhood			
NN(1)	12	0.303	160
NN(2)	14	0.084	160
NN(3)	14	0.057	160
NN(4)	14	0.067	160
NN(5)	14	0.069	160
Radius caliper			
0.01	11	0.459	51
0.0025	11	0.030	117
0.005	12	0.110	117
Kernel			
band width 0.01	11	0.459	51
band width 0.0025	14	0.038	117
band width 0.005	12	0.061	117

*Number of covariates with no statistically significant mean difference between matched samples of program participants and non-participants

Table A.3. Rosenbaum sensitivity analysis

Program-participation odd ratio (Γ)	Upper bound p-value from Wilcoxon sign-rank test
1	0.017335
1.1	0.026368
1.2	0.037451
1.3	0.050448
1.4	0.065171
1.5	0.081406
1.6	0.098931
1.7	0.11753
1.8	0.136995
1.9	0.157137
2	0.177784

B. FIRST STAGE REGRESSIONS FOR IV METHODS

Table B.1. First-Stage Regression with all Exclusion Restrictions

VARIABLES	Model Via	Model Vib	Model Vic	Model Vid
menja	0.657***	0.608***	0.650***	0.596***
	(0.0936)	(0.108)	(0.0926)	(0.106)
menjadensity	-0.651**	-0.606**	-0.631**	-0.575**
	(0.235)	(0.234)	(0.236)	(0.229)
menghhsz	0.00522*	0.00486*	0.00514*	0.00452*
	(0.00241)	(0.00237)	(0.00243)	(0.00229)
othpartcp	0.0448**			0.0612***
	(0.0192)			(0.0183)
dstpfm		-0.000803		-0.000805
		(0.000455)		(0.000472)
othfrst			-0.0531**	-0.0407
			(0.0234)	(0.0274)
offrma	-0.0201	-0.0229	-0.0137	-0.0252
	(0.0191)	(0.0164)	(0.0170)	(0.0164)
lndsza	0.0100*	0.0137**	0.0117*	0.0130**
	(0.00536)	(0.00590)	(0.00544)	(0.00551)
agea	9.67e-05	-0.000163	1.87e-05	-0.000138
	(0.000647)	(0.000638)	(0.000692)	(0.000615)
hhsize	0.000479	-1.00e-05	0.00182	-0.000933
	(0.00582)	(0.00702)	(0.00609)	(0.00672)
wealth	-0.0272*	-0.0171	-0.0259*	-0.0152
	(0.0134)	(0.0186)	(0.0139)	(0.0198)
hhdstwnmin	-0.000518	-0.000514	-0.000461	-0.000434
	(0.000431)	(0.000412)	(0.000409)	(0.000391)
hhdstroadmin	0.000581	0.000752	0.000563	0.000847
	(0.000722)	(0.000682)	(0.000671)	(0.000656)
tlua	-0.00505	-0.00456	-0.00531	-0.00406
	(0.00276)	(0.00316)	(0.00327)	(0.00337)
edu_max	-0.000276	-0.000160	-0.000377	4.94e-05
	(0.00345)	(0.00368)	(0.00348)	(0.00367)
edmiss	-0.00593	-0.00645	-0.00916	-0.00955
	(0.0282)	(0.0281)	(0.0304)	(0.0305)
fc	-0.444***	-0.464***	-0.436***	-0.458***
	(0.108)	(0.112)	(0.107)	(0.111)
vpopsze	0.000255***	0.000245***	0.000255***	0.000243***
	(4.39e-05)	(4.65e-05)	(4.32e-05)	(4.54e-05)
pop_miss	1.089***	1.019***	1.093***	1.009***
	(0.198)	(0.209)	(0.197)	(0.207)
Constant	-0.313**	-0.206	-0.301*	-0.190
	(0.137)	(0.189)	(0.138)	(0.190)
Observations	370	370	369	369
R-squared	0.839	0.844	0.840	0.846

Robust (clustered by village) standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1