Optimizing Voluntary Deforestation Policy in the Face of Adverse Selection and Costly Transfers

Arthur van Benthem and Suzi Kerr

Motu Working Paper 10-04
Motu Economic and Public Policy Research

August 2010

Author contact details

Arthur van Benthem
Stanford University, Department of Economics
arthurvb@stanford.edu

Suzi Kerr
Motu Economic and Public Policy Research
suzi.kerr@motu.org.nz

Acknowledgements

The authors would like to thank Adam Millard-Ball, the members of the PESD Offsets Working Group at Stanford University and several participants at the AIMES open science conference in Edinburgh 2010 for helpful comments and suggestions as well as Hugh McDonald for editing help.

Motu Economic and Public Policy Research

PO Box 24390
Wellington
New Zealand

Email	info@motu.org.nz
Telephone	+64 4 9394250
Website	www.motu.org.nz

ISSN 1176-2667 (Print), ISSN 1177-9047 (Online).

Abstract
As part of international climate change policy, voluntary opt-in programs to reduce emissions in unregulated sectors or countries have spurred considerable discussion. Since any regulator will make errors in predicting baselines, adverse selection will reduce efficiency since participants will self-select into the program. In contrast, pure subsidies lead to full participation but require large financial transfers; this is a particular challenge across countries. A global social planner facing costless transfers would choose such a subsidy to maximize efficiency. However, any actual policy needs to be individually rational for both the buying (industrialized) and selling (developing) country. We present a simple model to analyze this trade-off between adverse selection and infra-marginal transfers. The model leads to the following findings. First, extending the scale of voluntary programs both improves efficiency and reduces transfers. Second, the set of individually rational and Pareto efficient policies typically features a combination of credit discounting and stringent assigned baselines which reduce efficiency. Third, if the industrialized countries can be persuaded to be more generous, the feasible policy set can come close to the globally efficient policy to avoid deforestation.

JEL codes
Q54, Q56

Keywords
Voluntary opt-in; adverse selection; deforestation; offsets; emissions trading; REDD

1. Motivation

Many reports (e.g. Stern, 2006) and key policy makers assert that avoiding deforestation is a key short-run climate mitigation option because of the apparently low abatement costs (Kindermann et al., 2008). Melillo et al. (2009) and Wise et al. (2009) both show that it is critically important to price carbon in forests, especially if there are positive incentives for biofuels. Current estimates of the forest carbon supply curve are based on either land use responses to commodity prices (e.g. Kerr et al., 2002), or on estimates of the opportunity cost of land. These approaches do not take into account the difficulty of designing effective policies to address deforestation in developing countries, where most deforestation occurs (e.g. Andam et al., 2008; Blackman et al., 2009; Pfaff et al., 2007). They assume the application of efficient price-based policies, yet actual price-based policies for climate mitigation in developing countries are still mostly limited to offset programs. Examples include the payments for ecosystem services program in Costa Rica and the Clean Development Mechanism (CDM), where actors are given credit for forest remaining above an estimated and assigned baseline, or for emission reductions below a baseline.

Offset programs have been shown to suffer from serious problems of spurious credits and low effectiveness as a result of adverse selection. Adverse selection is caused by a combination of two factors: a *voluntary* element (i.e., agents can choose whether or not to opt in to the program) and *asymmetric information* about the baseline (i.e., the agents know more about their true baseline than the regulator).[1]

Several designs have been proposed for an international program to reduce deforestation (most recently referred to as REDD – reducing emissions from deforestation and degradation) and some are beginning to be implemented on a wider scale – notably Norway's innovative contract with Guyana.[2] All proposed policies have elements of offsets in their design and face a tradeoff between efficiency and the desire of the funders of such programs to get the best value for the money that they spend. This tradeoff exists because more generous or expensive programs (higher payments, with more favorable baselines) are more efficient.[3]

---

[1] Montero (2000), Fischer (2005) and Arguedas and van Soest (2009) establish theoretical results. Montero (1999) gives the first empirical evidence in the case of the US acid rain program. He and Morse (2010) and Millard-Ball (2010) explore similar issues in the energy sector for the CDM and for sectoral transportation caps respectively.

[2] See Chomitz (2007), Plantinga and Richards (2008) and Richards and Andersson (2000) for discussion of the challenges.

[3] Several studies provide evidence on the efficiency effects of adverse selection in the context of Costa Rican deforestation (Kerr et al., 2004; Robalino et al., 2008; Sanchez et al., 2007). Busch et al. (2009) focus on the global efficiency effects of different baselines (reference levels) in a deforestation program.

To the best of our knowledge, no one has previously formally analyzed three key options to reduce this tradeoff in a realistic policy context. In addition, while previous research has focused on designing mechanisms to reduce deforestation in developing countries in the presence of asymmetric information, the policies of the industrialized countries that buy credits or provide funds are often considered unconstrained (Strand, 1997). This paper explicitly considers welfare in both developing and industrialized countries, and highlights the limitations that voluntary participation by industrialized countries places on the set of implementable policies. Specifically, the paper explores the role of project scale, discounting (paying less than the full value for reductions) and baseline manipulation (making baselines either more or less stringent than business as usual) in a formal model. We use a microeconomic model of land use and mechanism design with a combination of analytical results and numerical simulations to show that (1) increasing the required scale of projects that can participate both improves efficiency and lowers transfers; (2) if industrialized countries are averse to transfers to developing countries they will optimally use a mixture of discounting and stringent baselines at the cost of some efficiency loss; and (3) if the industrialized countries can be persuaded to be more generous it will be easier to create an efficient policy to avoid deforestation. Our presentation focuses on deforestation but the results are equally applicable to any other mitigation options in developing countries, as well as wider applications of voluntary offset programs.

Our paper can be interpreted as an analysis of either adding avoided deforestation to a broader cap and trade market, or as an international fund used to pay for avoided deforestation to supplement a separate cap on other emissions. Both programs involve a baseline level of forest and provide rewards relative to that. In a cap and trade market these rewards would be credits valued at the market price, whereas in the fund these rewards would be dollars. In both programs industrialized countries pay for the reductions that are achieved in developing countries. These two approaches are equivalent under the following assumptions. First, the rewards must be the same per unit of avoided deforestation. To set fund payouts that meet this assumption requires that the aggregate marginal cost functions of the forest landowners are known so that the market price in the cap and trade system can be predicted accurately. Second, the cap and trade market emissions cap and the level of the fund can be adjusted so that regardless of which approach is used, both the global environmental outcome and the permit price are identical. That is, the fund level would need to be set such that the environmental gains it achieved were equal to the difference in environmental gains between the environmental cap of the larger broad cap and trade system (including avoided deforestation) and the original cap and trade market (excluding avoided deforestation).[4]

Three criteria that are of direct relevance to policy makers are used to assess the impacts of different conditions and policies: efficiency; the level of avoided deforestation and the payments per hectare of avoided deforestation. Efficiency is determined by whether land goes to its optimal use – land that yields high agricultural or timber returns should be cleared; land with returns lower than the positive environmental externalities from the forest should not. While policy focuses only on climate related externalities, we might also be concerned about avoided deforestation as an end in itself; this is the motivation for the second criterion. The final criterion is concerned with the value that industrialized countries get for the money they transfer to developing countries.

The remainder of this paper is organized as follows. Section 2 presents a simple model of voluntary deforestation policy that operates first at the level of individual plots, and then for larger scales. This demonstrates the trade-off between efficiency loss from adverse selection and the level of transfers, and analyzes how the three policy criteria are affected by the shapes of the distributions of land returns and observation errors. Section 3 discusses how the potential objectives are affected by three different *policy choices*: increasing the scale required for participation, changing the carbon payment (equivalent to "discounting credits") and changing the generosity of the assigned baseline. Section 4 provides a framework to make optimal trade-offs among objectives while taking into account the divergent interests of developing and industrialized countries and numerically analyzes the optimal policy. Section 5 concludes and summarizes the main policy implications.

2. A Simple Model of Voluntary Opt-In

2.1. Efficient subsidies versus baselines with adverse selection

Consider a continuum of plots of forested land, indexed by i. Decisions on each plot are independent. Landowners decide to either clear fully or keep the forest. Landowners will clear

[4] Suppose industrialized countries (ICs) have a joint emissions cap that requires them to undertake abatement of A. Total abatement cost (*TAC*) is the integral under the IC marginal abatement cost curve up to A. The market price of pollution equals p_c. ICs could use the fund to achieve n further units of abatement (and pay for m infra-marginal, or "spurious", units). Total global abatement would be $A + n$.
Analogous to the fund, ICs could purchase $n + m$ offsets from developing countries (DCs). This, however, would not be a fair comparison. Under the fund, the global abatement equals $A + n$. Using offsets, global abatement will be $A - m$. The environmental outcome is worse than without offsets (and p_c would be lower). To correct this, ICs must increase their joint abatement target to $A + n + m$. This ensures that, after $n + m$ offsets are purchased from DCs, the IC mitigation effort is back at A and the pollution price at p_c. Global abatement is now also $A + n$.

their forest if the net return from deforesting r_i (agricultural plus timber revenues minus clearing costs) exceeds any payment p_c to maintain forest. Landowner i knows r_i with certainty. The marginal environmental externality from deforestation is defined as δ.[5] Returns r_i are distributed across i with density f_r.

The simplest policy would be to offer a subsidy equal to p_c per plot that remains forested, where $p_c = \delta$. All landowners with $r_i \leq p_c$ will accept the subsidy and not deforest but only landowners with $0 \leq r_i \leq p_c$ will actually change their behavior; landowners with $r_i > p_c$ will (efficiently) deforest. The change in economic surplus ΔS_{eff} from this efficient policy relative to no policy equals

$$Efficiency\ gain = \Delta S_{eff} = \int_0^{p_c} (p_c - r)\, f_r(r) dr \tag{1}$$

This achieves efficient deforestation but requires a large transfer of resources

$$Total\ transfer = TT = p_c Pr(r \leq p_c) = p_c \int_{-\infty}^{p_c} f_r(r)\, dr \tag{2}$$

The total amount of avoided deforestation is

$$Avoided\ deforestation\ = AD = \int_0^{p_c} f_r(r) dr \tag{3}$$

The payment for each unit of deforestation (*PAD*) is the ratio TT/AD which will be very high if many plots of land have negative returns and so would not have been cleared even without the subsidy.

To avoid large transfers, a second policy option is a voluntary deforestation program that will pay participants an amount p_c for each hectare of forest exceeding an assigned baseline.[6] Landowners know their true forest baselines BL_i:

$$BL_i = \begin{cases} 1 & if\ r_i \leq 0 \\ 0 & if\ r_i > 0 \end{cases} \tag{4}$$

If the regulator observes r_i, the efficient solution is achieved by assigning each landowner i the true baseline $BL_i(r_i)$. If $BL_i = 1$ (no deforestation), no payment will be made and the forest will remain intact. If $BL_i = 0$ (full deforestation) and $0 \leq r_i \leq p_c$, the landowner will opt in and

[5] For the purposes of this paper this includes only carbon emissions. More generally it could also include non-climate externalities, such as loss of biodiversity.

[6] If it were practically feasible, a policy that sets $p_c = r_i$ would reduce transfers even further. In a recent paper, Mason and Plantinga (2010) describe a model in which the regulator has the option to provide landowners with a menu of two-part contracts, which consist of a lump-sum payment from the landowner to the regulator and a "per unit of forest" back to the landowner. Under certain conditions, these are type-revealing, where an ex-ante unobserved "type" corresponds to a marginal opportunity cost curve of keeping a fraction of the land forested. A similar approach to maximize the benefits to the developed country funders in an environmental transfer program was developed in Kerr, 1995. Our model does not consider this option. Instead, we focus on single baselines and show that increased scale can both enhance efficiency and dramatically reduce transfers.

choose not to deforest. If $BL_i = 0$ and $r_i > p_c$, the landowner will deforest and forego the payment p_c. If $p_c = \delta$, the remaining deforestation is *efficient.* Efficiency and avoided deforestation are the same as in (1) and (3) but the total transfer is lower by the amount in (5) and hence the payment per unit of deforestation is lower. This policy dominates the subsidy if transfers are costly.

$$Reduction\ in\ TT\ relative\ to\ subsidy\ = p_c \int_{-\infty}^{0} f_r(r)\, dr \tag{5}$$

In practice, however, the regulator cannot observe r_i, but instead observes $\hat{r}_i = r_i + \varepsilon_i$. The observation error ε has density $f_\varepsilon \sim (0, \sigma_\varepsilon)$, is assumed to be symmetric around 0 and independent of f_r. The *predicted* baselines are

$$\widehat{BL}_i = \begin{cases} 1 & if\ r\hat{}_i \leq 0 \\ 0 & if\ r\hat{}_i > 0 \end{cases} \tag{6}$$

What happens if the government assigns baseline $\widehat{BL}_i$? When $(r_i > 0, \hat{r}_i > 0)$ or $(r_i \leq 0, \hat{r}_i \leq 0)$, the assigned baseline coincides with the true baseline. The landowner will make the socially efficient decision. However, if $(r_i > 0, \hat{r}_i \leq 0)$, the assigned baseline is 1 but the true baseline is 0. The landowner would have deforested the plot in the true baseline, but gets assigned an unfavorable "no deforestation" baseline. Hence, the landowner will not participate in the scheme. This leads to an efficiency loss if $0 \leq r_i \leq p_c = \delta$, since the landowner will now deforest while he would not have done so had his baseline been correctly assigned and he had participated in the scheme. Relative to the efficient outcome in (1) the efficiency loss caused by adverse selection equals

$$\int_0^{p_c} (p_c - r) \left(\int_{-\infty}^{-r} f_\varepsilon(\varepsilon) d\varepsilon \right) f_r(r) dr \tag{7}$$

The amount of avoided deforestation will fall by

$$Pr(0 \leq r \leq p_c, \hat{r} \leq\ 0) = \int_0^{p_c} \left(\int_{-\infty}^{-r} f_\varepsilon(\varepsilon) d\varepsilon \right) f_r(r) dr \tag{8}$$

Finally, consider the case where $(r_i \leq 0, \hat{r}_i > 0)$. These landowners would have kept their forest, but now get assigned a full deforestation baseline. This will not affect their behavior, but it implies an additional infra-marginal transfer p_c. The total transfer (*TT*) is now lower than the subsidy amount (2) but higher than in the full information voluntary program. Total transfers (*TT*) are given by the sum of *marginal transfers* (*MT*) and *infra-marginal transfers* (*IT*). The former are the payments made to landowners that change their decision to not deforesting as a result of the policy. The latter are payments to landowners that would not have deforested without the policy, but get assigned a favorable full deforestation baseline and will therefore opt in.[7]

$$TT = MT + IT$$
$$= p_c \int_0^{p_c} \left(\int_{-r}^{\infty} f_\varepsilon(\varepsilon) d\varepsilon \right) f_r(r) dr + p_c \int_{-\infty}^{0} \left(\int_{-r}^{\infty} f_\varepsilon(\varepsilon) d\varepsilon \right) f_r(r)\, dr \quad (9)$$

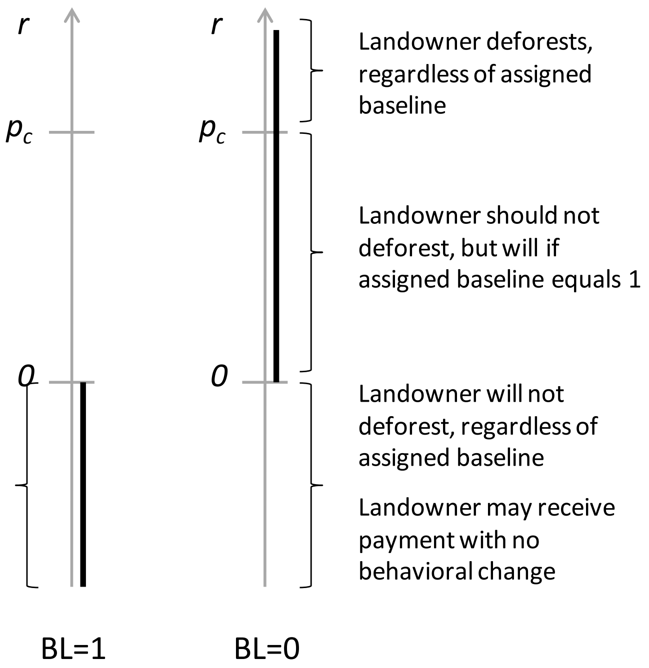
Because the amount of avoided deforestation is reduced relative to both the subsidy and the full information voluntary program while the total transfers lie between the two, the effect of adverse selection on payment per unit of avoided deforestation is theoretically ambiguous relative to the subsidy but clearly higher relative to the full information voluntary program.

To obtain intuition for this ambiguity, we use the decomposition in (9) to write *PAD* as

$$p_c \left(1 + \frac{\int_{-\infty}^{0} \left(\int_{-r}^{\infty} f_\varepsilon(\varepsilon) d\varepsilon \right) f_r(r)\, dr}{\int_0^{p_c} \left(\int_{-r}^{\infty} f_\varepsilon(\varepsilon) d\varepsilon \right) f_r(r) dr} \right) = p_c \left(1 + \frac{IMCF}{AD} \right) \quad (10)$$

where *IMCF* denotes the amount of infra-marginally credited. Moving from a subsidy to a voluntary program reduces *IMCF* but also lowers *AD*. For most realistic distributions the reduction in *IMCF* is larger than the reduction in *AD*, so *PAD* would fall. The cases described above are summarized in Figure 1.

Figure 1. Adverse selection causes efficiency loss in the range $0 \leq r_i \leq p_c = \delta$. It increases transfers in the range $r_i \leq 0$. Both are caused by assigning landowners in this range an incorrect baseline.



[7] In a cap and trade program, infra-marginal transfers would be spurious or non-additional credits.

2.2. The impacts of observation error and returns distributions on policy objectives

The tradeoff between efficiency and transfers depends on the distributions of observation errors and returns. We now analyze the impact of each of these on our three policy objectives: economic efficiency, avoided deforestation (*AD*) and payments per hectare of avoided deforestation (*PAD*).

2.2.1. Effects of observation error variance

Equation (7) shows that any change in $f_\varepsilon(\varepsilon)$ that increases the probability mass in the range [-∞, -*r*], where $0 \leq r \leq p_c$, will increase the efficiency loss from adverse selection (assuming $p_c = \delta$) and decrease avoided deforestation. A mean preserving spread such that $F_\varepsilon^{'}(x) \geq F_\varepsilon(x)\ \forall x < 0$ is sufficient. If the distribution of errors is normal, an increase in variance will generate such a mean preserving spread.

Under the same assumptions, *PAD* will increase. This follows from (10). Any increase in $f_\varepsilon(\varepsilon)$ that increases the probability mass in the range [-*r*, ∞], where $r \leq 0$, such as an increase in σ_ε in a normal distribution, will increase *IMCF*. Since *AD* decreases, *PAD* rises. More landowners with $r_i < 0$ will now get assigned $\widehat{BL}_i = 0$ and receive the payment p_c, but they do not provide *additional* deforestation.

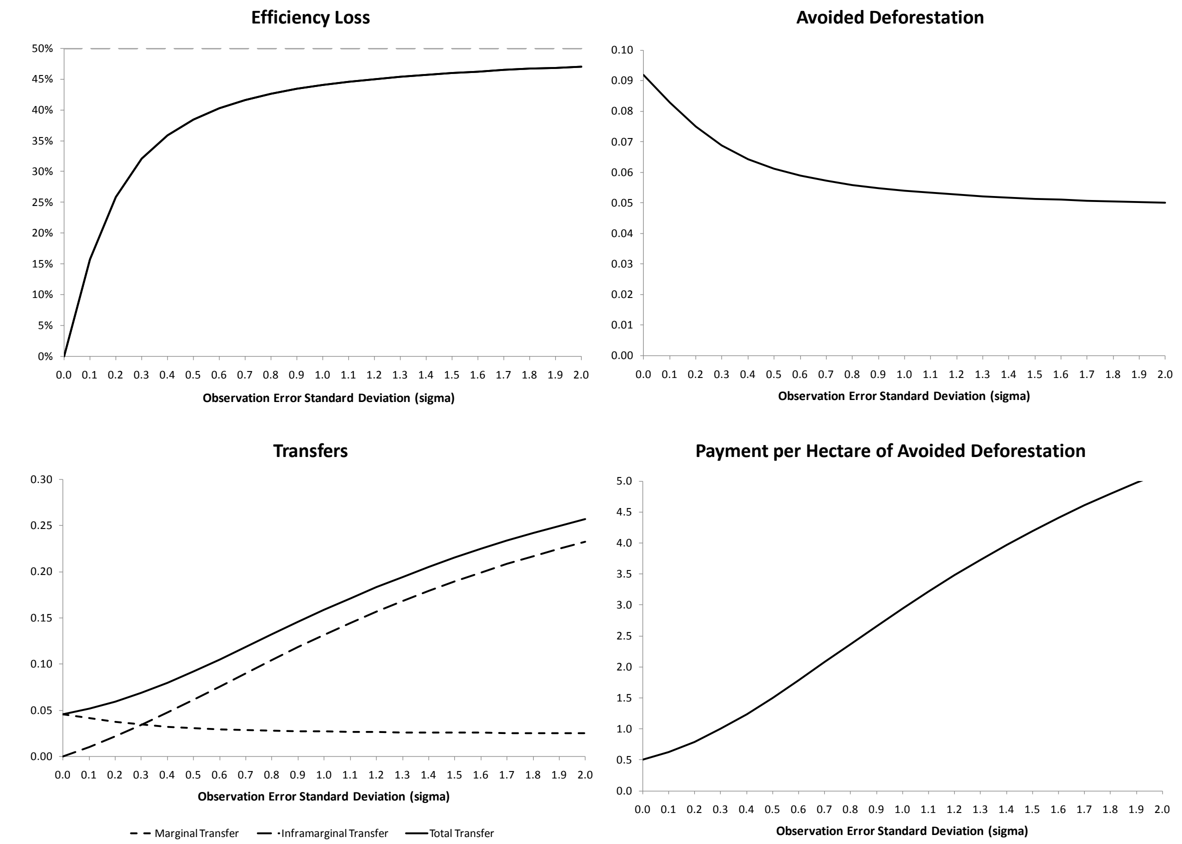
Numerical illustration.

To provide more intuition for the results with realistic distributions, we now assume a parametric form for the distribution of net agricultural returns $f_r(r)$ on forested land and the baseline prediction error $f_\varepsilon(\varepsilon)$. In the remainder of this paper, we will focus mostly on return distributions $f_r(r)$ for which $F_r(0) > 0.5$ and that are downward sloping at 0. The first assumption reflects that, at any point in time, landowners have previously chosen not to clear the remaining forest so only land on which relative returns have recently risen will still be forested but be at risk of clearing. The second assumption implies that there is a higher probability mass for returns just below zero than for returns just above zero, which intensifies the tradeoff between efficiency and reducing transfers and, in particular, infra-marginal rewards.

With no shocks, all land with positive returns would be cleared without any policy while no land with negative returns will be cleared. Hence, there will be positive probability mass below zero and no mass above zero and the assumption trivially holds. Deforestation occurs because the returns distribution shifts over time. If this shift, driven by, for example, technology and local infrastructure change, has both a common and an idiosyncratic (e.g. normal unbiased shock to each plot) element we would still expect the second assumption to hold.[8] The density above zero will tend to be lower than below zero, since the tail of the normal distribution implies a negative slope.

We consider $f_r(r) \sim N(-1,1)$, $f_\varepsilon(\varepsilon) \sim N(0, \sigma_\varepsilon)$ and $p_c = \delta = 0.5$ as our central case. Figure 2 plots the various policy objectives as a function of the standard deviation of the observation error σ_ε: the efficiency loss from adverse selection (7) relative to potential efficiency (1), *AD*, transfers and *PAD*.

Figure 2. Efficiency loss, *AD*, transfers and *PAD* as a function of observation error standard deviation σ_ε ($p_c = \delta = 0.5$).



Naturally, the efficiency loss is 0 if the observation error standard deviation $\sigma_\varepsilon = 0$. The efficiency loss is increasing in σ_ε. As σ_ε grows large the assignment of baselines becomes random. Participation, efficiency and avoided deforestation all fall toward 50% of their maxima at $\sigma_\varepsilon = 0$. Figure 1 shows that efficiency loss and avoided deforestation only result from landowners with $0 \leq r_i \leq \delta$. These will make the inefficient decision to deforest if and only if

[8] The common shock will generate a probability mass of forested land above zero return up to the size of the shock; the idiosyncratic shock will also move some land to higher returns (and some to lower) leaving lower probability mass in the upper tail of the returns distribution.

they get assigned $\widehat{BL}_i = 1$, which happens with probability approaching 0.5 as σ_ε increases. As σ_ε increases, infra-marginal transfers rise. Combined, the fall in *AD* and rise in transfers have dramatic implications for *PAD*; it quickly rises from the efficient value of 0.5 (the environmental externality δ). Even for a fairly modest σ_ε of 0.3, *PAD* doubles.

This section has shown that a mean preserving spread that increases the tails of the observation error distribution (which in a normal distribution would be implied by an increased variance) unambiguously has (weakly) negative effects on all three policy objectives. Any improvement in our ability to observe returns, or equivalently predict deforestation, would reduce the tradeoff between efficiency and transfers. This is a challenge for science and economics.

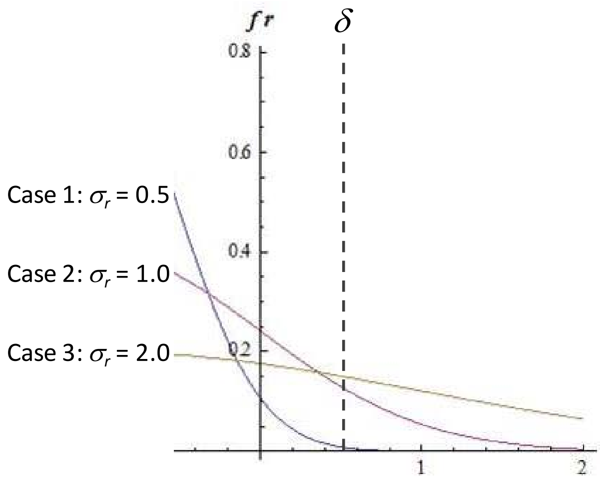
2.3. The impact of different marginal costs of avoiding deforestation on policy objectives

How do the policy objectives depend on the marginal abatement cost of avoided deforestation? In this model, abatement costs are represented by the foregone net return from deforestation *r* and the marginal abatement cost curve depends on the distribution $f_r(r)$.

The efficiency gain relative to no policy (1) depends on $f_r(r)$ through two effects. First, a higher probability mass of returns between [0, δ] increases the efficient level of *AD*. Second, a higher probability mass of very small positive returns between [0, $\varepsilon << \delta$] relative to returns between [δ - ε, δ] increases efficiency. Therefore, the first condition is not sufficient for an overall efficiency gain.

Figure 3 illustrates this for three different normal returns distributions. As the standard deviation σ_r grows from 0.5 to 1.0, f_r increases for all *r* between [0, δ]. This increases the efficiency gain of the policy. However, when σ_r increases further from 1.0 to 2.0, f_r increases for *r* close to δ, but decreases for small *r*. Hence, the effect on efficiency is ambiguous.

Figure 3. Returns distributions with μ_r = -1.0, σ_r = 0.5, 1.0 and 2.0, and σ_ε = 0.5.



Proposition 1. If the return distribution shifts from f to f' such that $f'_r(r) > f_r(r)$ for all r such that $0 \leq r \leq p_c = \delta$, $\Delta S(f') > \Delta S(f)$. If $\int_0^{p_c} f_r'(r)\,dr > \int_0^{p_c} f_r(r)\,dr$, then $AD(f') > AD(f)$.

Proof. Since $f_r'(r) > f_r(r)\ \forall r \in [0, p_c]$,

$AD(f') = \int_0^{p_c}\left(\int_{-r}^{\infty} f_\varepsilon(\varepsilon)d\varepsilon\right) f_r'(r)dr > \int_0^{p_c}\left(\int_{-r}^{\infty} f_\varepsilon(\varepsilon)d\varepsilon\right) f_r(r)dr = AD(f)$. Since $p_c - r \geq 0$,

$\Delta S(f') = \int_0^{p_c}(p_c - r)\left(\int_{-r}^{\infty} f_\varepsilon(\varepsilon)d\varepsilon\right) f_r'(r)dr \geq \int_0^{p_c}(p_c - r)\left(\int_{-r}^{\infty} f_\varepsilon(\varepsilon)d\varepsilon\right) f_r(r)dr = \Delta S(f)$.

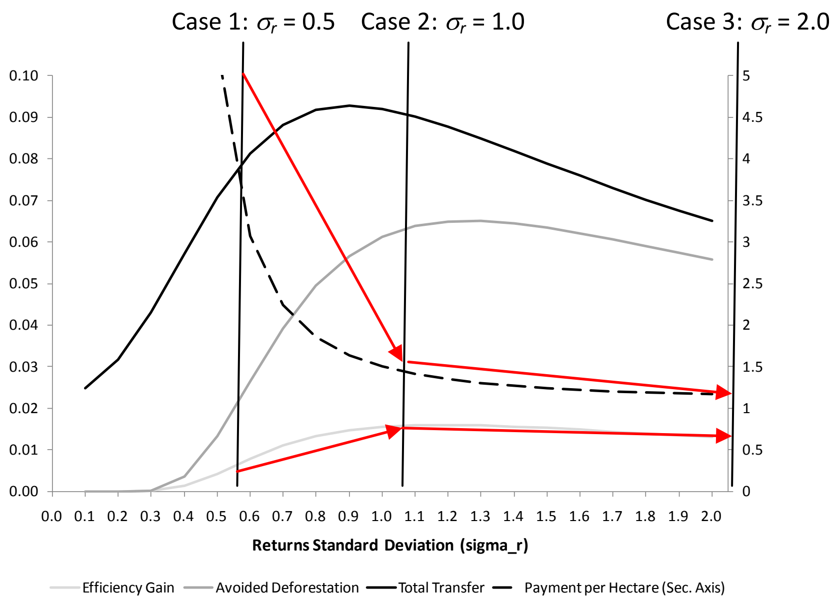
The second statement is trivial.

Proposition 1 shows that stronger assumptions on f_r and f_r' are needed to ensure an increase in efficiency than an increase in *AD*: an increase in the probability mass between [0, δ] is sufficient for *AD* to increase, but not to guarantee increased efficiency. With observation errors, a change in the returns distribution also affects the likelihood of infra-marginal payments: a less negatively (more positively) sloped distribution around zero yields lower infra-marginal payments. The combined effects on *AD* and infra-marginal payments determine the effect on *PAD*.

Numerical illustration.

We now illustrate these effects using our previous numerical example with $f_r(r) \sim N(-1,\sigma_r)$, $f_\varepsilon(\varepsilon) \sim N(0,\sigma_\varepsilon)$, $\sigma_\varepsilon = 0.5$, $p_c = \delta = 0.5$ and three different σ_r which alter the relevant part of $f_r(r)$.

Figure 4. Impact of changing $f_r(r)$ on the policy objectives, with $p_c = \delta = 0.5$.



Moving from case 1 to case 2 unambiguously raises efficiency and lowers *PAD*. This follows from Proposition 1, since $f'_r(r) > f_r(r)$ for all r between $[0, \delta]$. It corresponds to a downward movement in the marginal cost curve. In contrast, moving from case 2 to case 3, efficiency falls slightly; *PAD* does also. The probability mass of returns between $[0, \delta]$ decreases slightly, limiting the potential efficiency gains and *AD*. The distribution becomes almost flat in the region $[0, \delta]$. This means that the density close to zero (where costs are low) falls relative to the density close to δ (where costs are high). The marginal cost curve moves in such a way that less deforestation is avoided at a higher average cost. This flatness however also means that the ratio of land with returns at risk of infra-marginal payments (r just below 0) to returns with potential efficiency gains (r between $[0, \delta]$), and hence transfers and *PAD*, are lower.

A shift in the returns distribution that implies consistently lower marginal abatement costs in the relevant price range and reduces the density of returns just below zero relative to those above zero in the relevant price range will improve efficiency and reduce *PAD*. This might potentially be achieved by policies complementary to the voluntary program that address information failures or non-carbon externalities and hence increase the attractiveness of keeping land forested (e.g. for tourism, or a sustainable form of selective logging).

3. The Impact of Policy Choices

3.1. Policy 1: increasing the project scale

A first policy to consider is to increase the scale of each project. So far, we have considered a small-scale policy in which landowners get assigned plot-specific baselines and can opt in separately with each individual plot. While some forest carbon programs in practice are indeed small-scale, other proposals feature baselines for larger areas (e.g., a region or a country).[9] Section 2 showed that observation errors in voluntary programs reduce efficiency and avoided deforestation and increase payments per hectare of forest saved. This section shows that increasing the required scale of each project in the program mitigates these adverse consequences.

3.1.1. A multiple-plot model

We now consider a single landowner (alternatively, a region or country) who controls *n* 1-hectare plots. Each plot *j* has a return from deforestation r_j. We initially assume that these returns are distributed i.i.d. over plots with density f_r. Without the program, the landowner will clear all plots for which the return r_j exceeds zero. Hence, the true baseline is

$$BL_n = \sum_{j=1}^{n} BL_j \quad where \quad BL_j = \begin{cases} 1 & if\ r_j \leq 0 \\ 0 & if\ r_j > 0 \end{cases} \tag{11}$$

The government observes each r_j with error ε_j: $\hat{r}_j = r_j + \varepsilon_j$. Assume that ε_j is i.i.d. across *j*. This means that $\hat{r}_j$ has a distribution with mean μ_r and variance $\sigma_r^2 + \sigma_\varepsilon^2$. The distribution of $\hat{r}_j$ is more dispersed than $f_r(r)$. The government could compute an unbiased prediction of the baseline $\widehat{BL}_n$ as the sum of the expectation of the random variables for the plot-specific baselines. From its point of view, the true baseline for a specific plot is a Bernoulli random variable with mean p_{1i} and variance $p_{1i}*(1 - p_{1i})$, where $p_{1i} = \Pr(r_j < 0 | \hat{r}_j) = \Pr(BL_j = 1 | \hat{r}_j)$.[10] Since these are non-identically but independently distributed across *j*, the central limit theorem yields that for $n \rightarrow \infty$

---

[9] The Costa Rican Payments for Ecosystems services program is an example of a small scale system. Norway's recent agreement with Guyana sets up a large scale system.

[10] Note that $p_{1i} \neq \Pr(\hat{r}_j < 0)$, except if $f_r(r)$ is symmetric around zero. If the government naively assumed that *r* and $\hat{r}$ have the same distribution, it would calculate $p_{1i} = F_\varepsilon(-\hat{r}_j)$ = (if f_ε is symmetric) $1 - F_\varepsilon(\hat{r}_j)$. This would lead to a biased estimate of the baseline.

$$\widehat{BL}_n = \sum_{j=1}^{n} \widehat{BL}_j \xrightarrow{d} N\left(\sum_{j=1}^{n} p_{1j}, \sum_{j=1}^{n} p_{1j}(1-p_{1j})\right) \quad (12)$$

$\widehat{BL}_n$ is a continuous baseline for all n plots.[11] The landowner must now decide whether or not to opt in with his entire forest area, or not participate.

3.1.2. Increasing scale and efficiency

This section presents numerical simulations that show that a larger scale improves efficiency for "reasonable" returns distributions. In general, however, increasing the project scale has countervailing impacts on efficiency. The overall effect is ambiguous. Figure 5 provides intuition.

Figure 5. Single versus multiple plot policy.

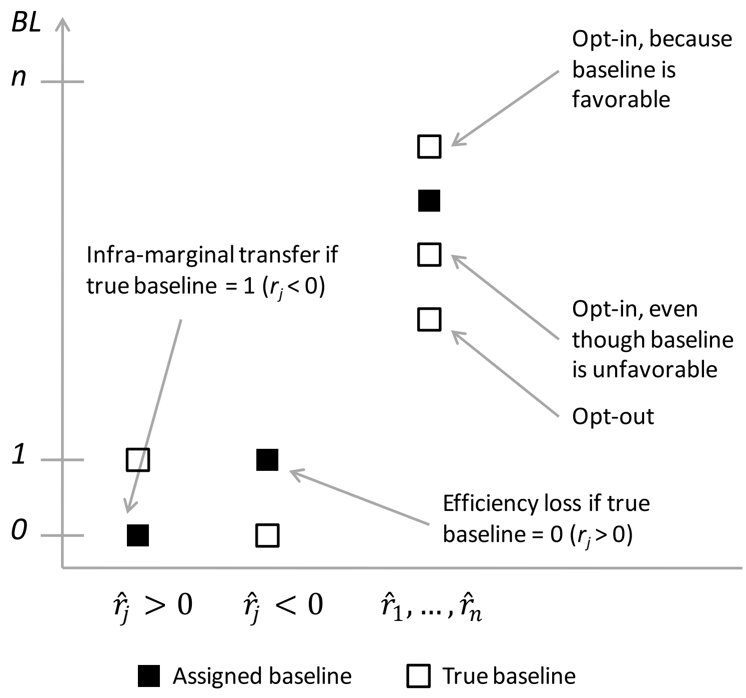


Figure 5 contrasts the single plot with the multiple plot case. In the single plot case, an inefficiency occurs when the true baseline is 0, but the government assigns a baseline of 1. In the multi-plot case, assigning a more favorable baseline $(\widehat{BL}_n < BL_n)$ will lead to guaranteed opt-in and infra-marginal payments. However, if $\widehat{BL}_n > BL_n$, the landowner has two options. If he opts in, he will clear all plots with returns exceeding $p_c = \delta_i$, but forego clearing plots with

[11] The baseline per plot will tend toward the true baseline and the observation error will tend toward zero per plot.

returns between 0 and p_c. Let n_{p_c} be the number of plots with $r < p_c$. Hence, opting in is favorable if and only if

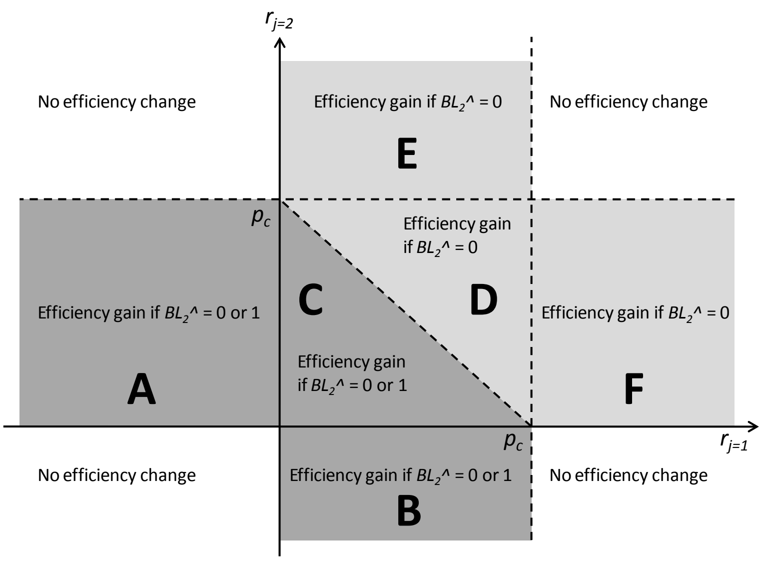
$$p_c * \left(n_{p_c} - \widehat{BL}_n\right) > \sum_{j\,|r_j \in [0,p_c]} r_j \tag{13}$$

Hence, for some assigned baselines $\widehat{BL}_n$, the landowner will still opt in, but for assigned baselines exceeding a threshold value, the landowner will opt out with all of his n plots. For this reason, the effect of increasing scale on efficiency is not a priori clear. There are cases in which scale increases efficiency: even if the baseline is too stringent, the landowner may still opt in *with all n plots*. Hence, all plots with returns between 0 and p_c will remain forested. This leads to a higher efficiency gain than plot-specific baselines, in which some plots with returns between 0 and p_c will get assigned a baseline equal to 1, and opt out.

However, in some cases the efficiency of the new system is lower than with plot-specific baselines. This happens when the baseline is so unfavorable that the landowners opt out *with all n plots*. In the single plot program, some critical plots with returns between 0 and p_c will receive a correct baseline and some deforestation will be efficiently avoided.

Comparing the overall efficiency gain from scaling up is not straightforward analytically. Figure 6 illustrates this for $n = 2$. With single-plot baselines, the expected efficiency gain relative to no policy is simply $2\Delta S$ with ΔS defined in (1) minus (7). Now consider a 2-plot baseline. The assigned baseline $\widehat{BL}_2 \in \{0,1,2\}$. Figure 6 shows under what circumstances this policy yields an efficiency gain relative to no policy.

Figure 6. Efficiency gains from a 2-plot baseline program.



The horizontal axis indicates the return of the first plot ($r_{j=1}$), and the vertical axis the (independently distributed) return of the second plot ($r_{j=2}$). We focus on the region in which there is a potential efficiency gain: either $r_{j=1}$, or $r_{j=2}$, or both returns are between 0 and p_c. In area A, the true baseline $BL_2 = 1$. With single-plot baselines, the efficiency gain would materialize only the assigned baseline for the second plot $\left(\widehat{BL}_{j=2}\right)$ was correct. However, the efficiency gain will now also occur if the $\widehat{BL}_{j=2} = 1$ (incorrect) but $\widehat{BL}_{j=1} = 0$ (incorrect), so that $\widehat{BL}_2 = 1$ (correct). Hence, errors in different directions may cancel each other. The same holds for area B. In area C, the full efficiency gain would materialize in a single-plot program only if $\widehat{BL}_{j=1} = \widehat{BL}_{j=2} = 0$. With a 2-plot program, the full gain is realized if the baseline is correct $\left(\widehat{BL}_2 = 0\right)$, but also if $\left(\widehat{BL}_2 = 1\right)$. This follows since (13) holds if $r_{j=1} + r_{j=2} < p_c$. This is an additional efficiency gain relative to a single-plot program. Hence, the dark grey shaded regions indicate combinations of returns for which a multi-plot program is more efficient than a single-plot program.

There is a flip side to this additional efficiency gain, represented by the light grey shaded regions. First, area D represents a region in which the full efficiency gain (relative to no policy) is realized *only* if $\widehat{BL}_2 = 0$. When $\widehat{BL}_2 = 1$, the landowner will opt out with all plots since (13) does not hold. In a single-plot program one of the plots may have been assigned $\widehat{BL}_j = 0$ and an efficiency gain (relative to no policy) would have been realized. Similar arguments hold for regions E and F.

Hence, to calculate if a multi-plot program constitutes an efficiency gain relative to a single-plot program, we need to compute the relative importance of the dark grey versus the light grey regions in Figure 6. Proposition 2 shows that the result is ambiguous and depends on the probability distributions f_r and f_ε.

Proposition 2. Even if f_r and f_ε are both symmetric and i.i.d. with mean 0, and f_r and f_ε are independent, 2-plot baselines are not always more efficient than single-plot baselines.

Proof. See Appendix 1.[12]

A proof for the efficiency of an *n*-plot program is analytically difficult, but the next section presents numerical simulations to investigate the implications for realistic distributions.

[12] Analogous to the proof of Proposition 2, 2-plot baselines do not always lead to higher *AD* than single-plot baselines. The impact of scale on *PAD* is also ambiguous.

3.1.3. Numerical simulations of increased scale

This section presents numerical simulations to illustrate the differences between a single-plot versus multiple plot program. Throughout this section, we assume a normal observation error distribution with $\sigma_\varepsilon = 0.5$, unbiased assigned baselines, and $p_c = \delta = 0.5$. The central case returns distribution is $f_r(r) \sim N(-1,1)$. We also consider alternative distributions.[13] In all cases, $f_\varepsilon(\varepsilon) \sim N(0,0.5)$. Table 1 demonstrates what happens to the policy objectives as n increases.

Table 1. The impact of increasing the required project scale: $f_r(r) \sim N(-1,1)$.

	1-plot	2-plot	10-plot	100-plot	Maximum potential $\sigma_\varepsilon = 0$
Δ efficiency	1.3331	1.4157	1.8248	2.4744	2.5319
AD	5.3387	5.5368	6.5951	8.9545	9.1848
TT	4.8817	4.8158	4.5139	4.5562	4.5924
PAD	0.9149	0.8702	0.6845	0.5089	0.5000
Opt in	4.88%	18.30%	58.42%	97.47%	100%

Notes: Each row is based on 10,000 random draws from the probability distributions. $\delta = p_c = 0.5$. $f_\varepsilon(\varepsilon) \sim N(0,0.5)$. *Δefficiency*, *AD* and *TT* are all normalized per 100 plots. Assigned baselines are unbiased.

Table 1 shows that increasing the project scale has dramatic consequences for its performance. For the central case, n-plot baselines increase efficiency and *AD* and reduce *PAD* as n increases. 100 plots are enough to approach the efficient solution. The reason is that the observation error *normalized per plot* decreases as n grows, and the probability of opt-in becomes very high (97.47%).[14] This high opt-in rate signals efficiency though Table 1 shows that most gains are achieved through the first 5% of plots that participate. Hence, scale mitigates adverse selection for the central case returns distribution. Appendix 2 demonstrates that, as discussed in Section 3.1.2, this result does not hold for all distributions.

We now test the robustness of these conclusions. An important assumption has been that $f_r(r)$ and $f_\varepsilon(\varepsilon)$ are i.i.d: both returns and observation errors are independent across plots. In reality, there may be a high degree of spatial correlation in both returns and errors. We introduce spatial correlation across plots in the following stylized way:

[13] In addition to the central case returns distribution, we consider $f_r(r) \sim N(0,1)$ and $f_r(r) \sim$ symmetric bimodal normal with modes -0.5 and 0.5, and standard deviation 0.1. Results for these are given in Appendix 2.

[14] In some realizations, landowners *efficiently* opt out: their return exceeds δ.

$$r_j = \rho_r r_{j-1} + u_r$$
$$\varepsilon_j = \rho_\varepsilon \varepsilon_{j-1} + u_\varepsilon \quad (14)$$

where u_r and u_ε are i.i.d. with variances σ_{ur}^2 and $\sigma_{u\varepsilon}^2$, such that $\sigma_r^2 = \sigma_{ur}^2 / (1 - \rho_r^2)$ and $\sigma_\varepsilon^2 = \sigma_{u\varepsilon}^2 / (1 - \rho_\varepsilon^2)$. Table 2 summarizes the main findings for the central case. Results for other distributions are given in Appendix 2.

Table 2. The impact of spatially correlated returns and observation errors: $n = 1$ and $n = 100$. $f_r(r) \sim N(-1,1)$.

	1-plot	100-plot	100-plot	100-plot	100-plot	$\sigma_\varepsilon = 0$
		$\rho_r = 0$, $\rho_\varepsilon = 0$	$\rho_r = 0.5$, $\rho_\varepsilon = 0.5$	$\rho_r = 0.9$, $\rho_\varepsilon = 0.9$	$\rho_r = 0.99$, $\rho_\varepsilon = 0.99$	
Δefficiency	1.3331	2.4744	2.4130	2.0204	1.5901	2.5319
AD	5.3387	8.9545	8.7245	7.2744	5.7233	9.1848
TT	4.8817	4.5562	4.5228	4.4151	4.5881	4.5924
PAD	0.9149	0.5089	0.5184	0.6070	0.8024	0.5000
Opt in	4.88%	97.47%	94.96%	78.40%	37.28%	100%

Notes: Each row is based on 10,000 random draws from the probability distributions $\delta = p_c = 0.5$. $f_\varepsilon(\varepsilon) \sim N(0,0.5)$. *Δefficiency*, *AD* and *TT* are all normalized per 100 plots. Assigned baselines are unbiased.

Table 2 shows that only very high spatial correlation ($\rho \geq 0.9$) undermines the effect of increasing scale. As the correlation across plots and errors increases, efficiency and *AD* decrease, and *PAD* increases. The intuition is that observation errors do not cancel out across plots, but are persistent. High spatial correlation reduces the probability of participation and therefore adversely impacts the policy objectives. Larger required project scale would mitigate this.

3.2. Policy 2: choosing a different payment per hectare

In the analysis above, the payment per hectare p_c was assumed to be equal to the marginal externality from deforestation δ. This section analyzes what happens to the three project objectives if we vary p_c. Reducing p_c is equivalent to the practice of "credit discounting" sometimes observed in practice in offset systems (Kollmuss et al., 2010). Under a system of discounting, fewer offset credits are awarded than the environmental gains represented by the difference between the baseline and the actual forest level. First, we discuss the impact of changing p_c in a single-plot model. Then, we analyze how these results change in a multiple-plot model using numerical simulations.

It is straightforward to see that, independent of scale any $p_c \neq \delta$ is less efficient if $\sigma_\varepsilon = 0$. All landowners get assigned the true baseline, and paying less than δ reduces efficiency, because landowners with average returns $((\underline{r} \,|\, 0 \leq r \leq \delta) > p_c)$ would opt out. Paying more than δ reduces efficiency because some landowners will opt in even though their private gains from deforestation exceed the full environmental cost. This is inefficient from an economic perspective.[15] In the single-plot model with full and symmetric information, the change in efficiency relative to a no policy case was given in (1). A simple application of Leibniz' Rule yields that efficiency is maximized when $p_c = \delta$. We will now investigate if this result changes with asymmetric information – i.e. when $\sigma_\varepsilon > 0$.

3.2.1. Different payments in the single-plot model

In the single-plot model, the introduction of observation error does not change the conclusion that the most efficient payment is $p_c = \delta$. The efficiency change relative to no policy equals

$$\Delta S(p_c) = \int_0^{p_c} (\delta - r)\left(\int_{-r}^{\infty} f_\varepsilon(\varepsilon) d\varepsilon\right) f_r(r) dr \tag{15}$$

Proposition 3. In the single-plot model, efficiency is maximized for $p_c = \delta$, regardless of $f_\varepsilon(\varepsilon)$.

Proof. Using Leibniz' Rule, the first order condition is given by $\frac{d(\Delta S(p_c))}{dp_c} = (\delta - p_c)\left(\int_{-p_c}^{\infty} f_\varepsilon(\varepsilon) d\varepsilon\right) f_r(p_c)$. Since $\int_{-p_c}^{\infty} f_\varepsilon(\varepsilon) d\varepsilon > 0$ for any $f_\varepsilon(\varepsilon)$ and $f_r(p_c) \geq 0$, efficiency is maximized when $p_c = \delta$.

We now investigate what happens to the other two project objectives as the payment p_c varies.

Proposition 4. Avoided deforestation is (weakly) increasing in p_c. *MT*, *IT* and *TT* are globally (weakly) increasing in p_c.

[15] This is not a statement that current payments should not be increased. If current payments are considerably below the true environmental externality, more deforestation should be avoided from an economic perspective.

Proof. $AD = \int_0^{p_c} \left(\int_{-r}^{\infty} f_\varepsilon(\varepsilon) d\varepsilon \right) f_r(r) dr$. The derivative of AD w.r.t. p_c is $\left(\int_{-p_c}^{\infty} f_\varepsilon(\varepsilon) d\varepsilon \right) f_r(p_c) \geq 0 \ \forall p_c$, proving the first statement. The derivative of MT (first term in (9)) w.r.t. p_c is $\int_0^{p_c} \left(\int_{-r}^{\infty} f_\varepsilon(\varepsilon) d\varepsilon \right) f_r(r) dr + p_c \left(\int_{-p_c}^{\infty} f_\varepsilon(\varepsilon) d\varepsilon \right) f_r(p_c) \geq 0 \ \forall p_c$. The derivative of IT (second term in (9)) w.r.t. p_c is $\int_{-\infty}^{0} \left(\int_{-r}^{\infty} f_\varepsilon(\varepsilon) d\varepsilon \right) f_r(r) \, dr \geq 0 \ \forall p_c$. Hence, the derivative of TT w.r.t. p_c is weakly greater than zero $\forall p_c$.

The effect of changing p_c on *PAD* is ambiguous. To get some intuition for why this is the case, consider (10). *IMCF* denotes the amount of infra-marginally credited forest (which is independent of p_c). Avoided deforestation *AD* is increasing in p_c. Since *AD* is bounded, very high values of p_c will lead to increasing *PAD*. For intermediate values of p_c, a small increase in p_c can either lead to almost no additional deforestation, or a large increase in avoided deforestation, depending on the specification of the return distribution $f_r(r)$. For instance, if $f_r(r) = 0$ for $r \in [0, \underline{p}]$, then *PAD* will be infinite for $p_c \leq \underline{p}$ and achieve a global minimum for some $p_c > \underline{p}$. Hence, *PAD* can either be increasing or decreasing in p_c.[16]

We conclude that, in the single-plot model, efficiency is maximized by paying $p_c = \delta$. Paying more leads to more *AD*, but requires higher transfers. The effect on *PAD* is ambiguous for low values of p_c, but eventually *PAD* must increase.

3.2.2. Different payments in the multi-plot model

In the multi-plot model, $p_c = \delta$ no longer unambiguously maximizes efficiency. The intuition is as follows. Raising p_c above δ has two countervailing effects on efficiency. First, it will increase the opt-in probability. This increases efficiency because it helps prevent deforestation of plots with returns below δ. Second, it causes certain forest to be *inefficiently* prevented from deforestation. The relative strength of these channels determines whether a higher p_c can be more efficient than $p_c = \delta$. A lower p_c will never increase efficiency, since it will both reduce opt-in and cause inefficient deforestation. The effects go in the same direction. Hence, $p_c \geq \delta$ maximizes efficiency in the multiple-plot model. *PAD* is ambiguous in p_c. Table 3 illustrates this when $\delta = 0.5$ and $n = 10$ or 100.

Table 3 shows that raising p_c above δ can increase efficiency. For $n = 10$, raising p_c above δ (to $p_c = 0.6$) slightly increases efficiency. Hence, efficiency is no longer maximized at $p_c = \delta$. However, when $n = 100$, the opt-in probability at $p_c = \delta$ is already almost efficient at 97.47%.

[16] Proof available from the authors on request.

Raising p_c to 0.6 increases opt-in only slightly to 98.93%. Hence, we find that the most efficient solution is sometimes achieved for $p_c > \delta$. This increased efficiency coincides with higher *PAD*, however.

Table 3. The impact of changing p_c, for $f_r(r) \sim N(-1,1)$, $\delta = 0.5$ for $n = 10$ and 100.

		carbon payments (p_c)						
	n	0.1	0.25	0.4	0.5	0.6	0.75	1.0
Δefficiency	10	0.5290	1.2046	1.6650	1.8248	1.8754	1.7787	1.2991
	100	0.6751	1.6996	2.3279	2.4744	2.4496	2.2171	1.5886
AD	10	1.1723	3.1523	5.2428	6.5951	7.8780	9.6013	11.7763
	100	1.4928	4.4394	7.3462	8.9545	10.2858	11.8285	13.5757
TT	10	0.4282	1.5186	3.1517	4.5139	6.0625	8.6547	13.3455
	100	0.2541	1.2716	3.0522	4.5562	6.2118	8.8767	13.5741
PAD	10	0.3655	0.4819	0.6012	0.6845	0.7696	0.9015	1.1333
	100	0.1703	0.2865	0.4155	0.5089	0.6039	0.7505	0.9999
Opt-in	10	33.59%	44.25%	53.26%	58.42%	62.83%	68.19%	74.29%
	100	63.39%	84.16%	94.43%	97.47%	98.93%	99.74%	99.97%

Notes: Each row is based on 10,000 random draws from the probability distributions. $\delta = p_c = 0.5$. $f_\varepsilon(\varepsilon) \sim N(0,0.5)$. *Δefficiency*, *AD* and *TT* are all normalized per 100 plots. Assigned baselines are unbiased.

In summary, we find that efficiency considerations never justify setting the payment p_c below the environmental damage δ, and setting $p_c > \delta$ can be justified if opt-in at $p_c = \delta$ is below 100%. Since increasing scale leads to full opt-in in the limit, $p_c = \delta$ always becomes the most efficient payment as *n* approaches infinity. There always exists a p' ($p' > \delta$) such that efficiency falls and transfers rise for any $p > p'$. In that price region, the outcome is unambiguously worse. For choices $p_c < \delta$, the tradeoff between efficiency and transfers remains.

3.3. Policy 3: changing the generosity of the assigned baseline

Another policy choice for the regulator is to set a baseline that is, in expectation, too high or too low. In other words, the government assigns the following baselines for plot *i*

$$\widehat{BL}_i = \begin{cases} 1 & if\ r^\wedge{}_i \leq r^* \\ 0 & if\ r^\wedge{}_i > r^* \end{cases} \tag{16}$$

where r^* is a specified return set by the government. The government, aware of adverse selection, may try to only pay landowners who are most likely to deforest in the baseline, by choosing $p_c > r^* > 0$. Assuming $p_c = \delta$, we analyze the impact of this policy change on the three criteria: efficiency, *AD* and *PAD*. To provide intuition, we first discuss the impact in the context of the single-plot model. Then, we present numerical simulations of the multiple-plot model.

3.3.1. Changing baselines in the single-plot model

Proposition 5. More generous baselines increase efficiency and *AD*, but require a higher *TT*.

Proof. The efficiency change relative to no policy equals $\Delta S(r^*) = \int_0^{p_c}(p_c - r)\left(\int_{r^*-r}^{\infty} f_\varepsilon(\varepsilon)d\varepsilon\right)f_r(r)dr$. By Leibniz' Rule, this expression is globally weakly decreasing in r^*. Hence, efficiency is maximized if $r^* \to -\infty$. *AD* is given by $AD(r^*) = Pr(0 \leq r \leq p_c, \hat{r} > r^*) = \int_0^{p_c}\left(\int_{r^*-r}^{\infty} f_\varepsilon(\varepsilon)d\varepsilon\right)f_r(r)dr$, which is globally weakly decreasing in r^*. Finally, *TT* is given by $TT(r^*) = p_c\int_{-\infty}^{p_c}\left(\int_{r^*-r}^{\infty} f_\varepsilon(\varepsilon)d\varepsilon\right)f_r(r)dr$, which is also globally weakly decreasing in r^*.

The fact that efficiency increases as $r^* \to -\infty$ is not surprising, since this is equivalent to assigning a no-forest baseline or a subsidy of p_c per hectare of forest standing. As discussed in Section 2, such a subsidy is indeed efficient but requires a large infra-marginal transfer.

Using (10) and making *IMCF* and *AD* functions of r^* we can see that the effect of r^* on *PAD* is ambiguous. *IMCF*, the amount of infra-marginally credited forest, is decreasing in r^*, but so is *AD*. The shape of *PAD* is dependent on the return distribution $f_r(r)$.

Numerical illustration.

Figure 7. Impact of baseline generosity on the project objectives for the central case.

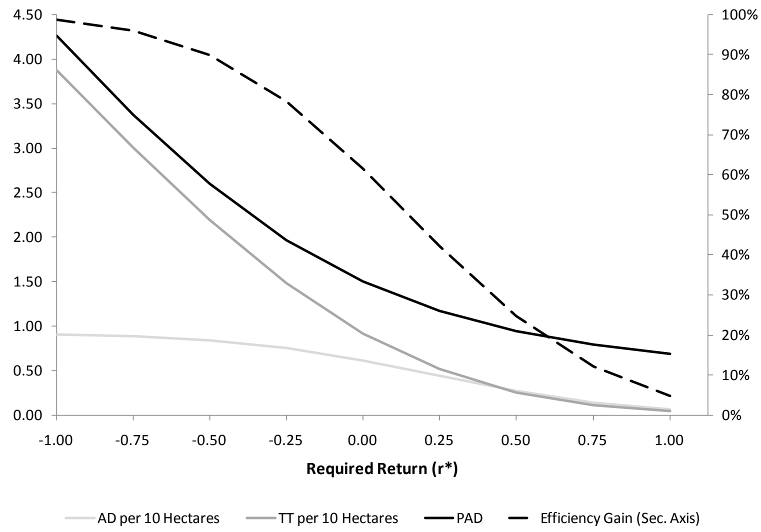


Figure 7 confirms that efficiency, *AD* and *TT* are decreasing in *r**. For "typical" returns distributions like our central case ($f_r(r)$~N(-1,1)), *PAD* is also decreasing in *r**. Hence, efficiency and *PAD* are generally conflicting policy aims for this policy option also: efficiency requires setting *r** low, while minimizing *PAD* requires setting *r** high.

3.3.2. Changing baselines in the multiple-plot model

The conclusions from the single-plot model also hold in the multiple-plot model. Table 4 illustrates the effect of assigning baselines that are too (un)favorable in expectation for the central case returns distribution.

Table 4. The impact of changing baseline generosity, for $f_r(r)$ ~ N(-1,1) and n = 100.

	expected assigned baseline ($\widehat{BL}_{100}$)					
	78	81	84 (unbiased)	87	90	93
Δefficiency	2.5264	2.5255	2.4744	2.0324	0.8671	0.0850
AD	9.1723	9.1526	8.9545	7.2768	3.0144	0.2798
TT	7.5860	6.0750	4.5562	2.7654	0.9043	0.0695
PAD	0.8271	0.6638	0.5089	0.3801	0.3001	0.2483
Opt-in	100.00%	99.85%	97.47%	79.64%	34.18%	3.54%

Notes: Each row is based on 10,000 random draws from the probability distributions. $\delta = p_c$ = 0.5. σ_ε = 0.5. n = 100.

In Table 4, the true baseline equals 84 (84 out of 100 plots will remain forested in absence of a policy). The table shows that increasing the baseline (i.e., making it less favorable) unambiguously reduces efficiency, *AD*, and opt-in probability, but also reduces *PAD*. This illustrates the conflicting policy objectives: achieving efficiency and avoiding deforestation comes at the expense of increased transfers for both single and multiple-plot programs.

This section has shown that only increasing project scale improves all objectives simultaneously (for "typical" returns distributions). Discounting credits and changing baseline generosity affect the objectives in opposing directions. This raises the question about how these objectives should be weighed against each other. In other words: what is the optimal deforestation policy? This question is addressed in the next section.

4. Choosing an Optimal Policy

The previous sections have illustrated that deforestation policies involve trade-offs between environmental outcomes, financial transfers and efficiency. This section provides a framework for how decision makers can optimally trade off these outcomes and then simulates the effects of different policy choices.

4.1. A framework for 'optimal' decision making

The optimal policy depends on whose interests are considered. A global social planner, who can require participation and faces no costs of transfers, could maximize efficiency and then redistribute the surplus to meet distributional (and political) objectives.[17] This is a useful benchmark, which we refer as the "globally efficient policy". However, in an international context, both the industrialized countries (ICs) that fund the policy and developing countries (DCs) that avoid deforestation must participate voluntarily. Our basic policy design ensures that the developing country participation constraint is met. We must now consider the participation constraint for the ICs. If we design a policy to maximize the surplus of the ICs subject to voluntary participation by DCs (so they more than meet their participation constraint) we will find a set of Pareto optimal policies.[18] As long as there is a global aggregate surplus from the policy, the stakeholders should in theory be able to negotiate a sharing rule for the surplus and all be better off with than without the policy.

[17] This is closer to a model of domestic regulatory design and implicitly assumes a utilitarian social welfare function.
[18] If we allowed for policies that induce full revelation of landowners' information we could potentially expand this set. We are not convinced that these are empirically feasible and even if they were, they seem likely to be difficult to implement internationally.

As discussed in Section 1, under assumptions we maintain in this paper, including avoided deforestation in a cap and trade system and rewarding it through an international fund are equivalent. However, when considering the cost of transferring resources between countries the systems could differ. ICs might be more likely to raise resources for a fund through distortionary taxation, while reductions under cap and trade might be an obligation placed directly on companies (implicitly a lump sum cost). However, given the politics of free allocation with cap and trade systems, the marginal resources for either could come from additional taxes or through reduced free allocation. From the DC point of view both approaches provide useful government revenue though they could be perceived differently. Thus without loss of generality we will present the optimal policy in terms of an international fund.

The fund is used to pay DCs that opt into the program. They receive p_c per hectare, which leads to n hectares of avoided deforestation and m units of infra-marginal forest receiving payments. They value each dollar received at its face value. DCs forego returns r_j on each of the n hectares of avoided deforestation. We assume that the global environmental gain δ is fully valued by ICs, and is not valued by DCs.[19] Payments are directly costly to ICs. In addition, transferring resources is associated with an additional cost or benefit of γ. If $\gamma = 0$, a one dollar transfer has a cost to ICs of exactly one dollar. Alternatively, $\gamma > 0$ refers to a situation in which distortionary taxes need to raise money for the fund or the IC has an aversion to paying for abatement in a DC relative to in their own country at the same cost. A third possibility is that $\gamma < 0$. Rich ICs may derive satisfaction from donating funds to poor DCs or may feel a sense of responsibility to cover a higher share of the costs of international climate mitigation.[20]

Under these assumptions, the costs and benefits to the different countries are

$$\begin{aligned} Gains\ to\ DCs\text{:}\ \ n\left(\frac{\int_0^{p_c}(p_c - r)\,f_r(r)dr}{\int_0^{p_c} f_r(r)\,dr}\right) + mp_c \\ = n(p_c - E[r|0 \le r \le p_c]) + mp_c \end{aligned} \tag{17}$$

[19] In reality, some of the environmental benefits accrue to DCs, whose governments may or may not value them. If they are valued by DC governments, it should lead to intervention even without an IC-induced policy. This shifts the baseline upwards. This would be an easy adjustment in our model, and would not affect the qualitative results.

[20] Note that from the perspective of the global social planner, $\gamma = 0$ implies costless transfers from ICs to DCs. This case corresponds to the "globally efficient policy", which is effectively a pure subsidy on forest standing. Even when transfers are costly to the DC ($\gamma > 0$), we assume that the global social planner could do costless lump-sum transfers without any distortionary effects. The case $\gamma < 0$ is more subtle. In principle, a global social planner would want to do an unlimited amount of redistribution from ICs to DCs, even beyond the full forest subsidy. To limit transfers, we must assume that γ is rising in the scale of transfer, possibly responding to the income difference between the groups or a limited "warm glow" effect of donations. We abstract from this issue and will continue to compare any feasible solutions to the same "globally efficient policy" as for $\gamma \ge 0$.

$$\mathit{Gains\ to\ ICs}:\quad \delta n - p_c(n+m)(1+\gamma) \tag{18}$$

$$\begin{gathered}\mathit{Net\ global\ gains\ with\ equal\ utility\ weights}: \\ n(\delta - E[r|0 \le r \le p_c]) - \gamma p_c(n+m)\end{gathered} \tag{19}$$

This specification of surplus combines the policy objectives discussed in previous sections: efficiency is captured by the first terms of the gains to DCs and ICs, *AD* corresponds to n, total transfers are included in the second term in the gains to ICs and m is equal to *IMCF* which affects *PAD*. Hence, this framework represents a method to balance the project objectives optimally.

As discussed in footnote 19, a global social planner will choose the efficient solution based on costless transfers ($\gamma = 0$) and then reallocate resources. The actual net global gains when transfers are costly/beneficial ($\gamma \neq 0$) are given in (19) and differ from the globally efficient policy. When the IC participation constraint (equation (18) ≥ 0) is added so that the gains to ICs must be non-negative, we can compute a set of *feasible* (individually rational) policies. Our model does not predict which feasible policy will be chosen, except that the policy must be Pareto efficient. The set of *feasible* policies does not necessarily contain the globally efficient policy. Equation (18) shows that if $m, \gamma > 0$, setting $p_c = \delta$ will never lead to IC participation.[21] Any such policy would imply negative gains to ICs. Two options could induce IC participation. First, credits could be discounted ($p_c < \delta$). Second, the baseline could be set stringent enough to achieve $m < 0$. Both will reduce efficiency. If γ is negative for less than some level of transfer, implying that ICs derive some satisfaction from transferring money to DCs, there will be less need to discount credits or require stringent baselines to meet the ICs' participation constraint and the loss of efficiency is reduced.

4.2. Numerical simulations of optimal policy choices

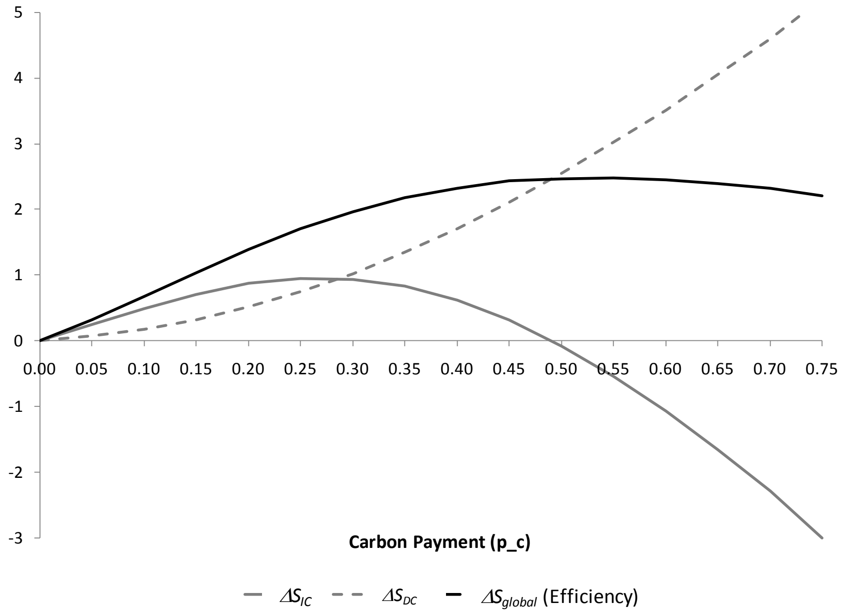
In this section, we use the numerical simulations to determine the optimal policy, both from the global planner's and the IC-DCs' perspectives. Initially we assume that $\gamma = 0$. Using our earlier notation, the IC participation constraint is now simply $\delta AD - TT \geq 0$. We assume that the policy will be a multiple-plot program ($n = 100$; subscript dropped below), since large scale

[21] If $\gamma = 0$, $p_c = \delta$ and $m = 0$ (e.g., if $\sigma_\varepsilon = 0$ and the assigned baseline is unbiased), ICs are indifferent about participation.

programs dominate small scale programs (Section 3). There are two policy levers: reducing p_c relative to δ and biasing downwards the assigned baseline $(\widehat{BL})$ relative to its expected value.

Section 3 presented simulation results in which these two policy levers were varied separately (Tables 3-4). Table 3 shows the impact of changing p_c, while $\widehat{BL}$ is unbiased and fixed at 84.[22] Figure 8 below shows the change in global surplus ΔS_{global} (=Δefficiency), as well as the change in surplus for ICs ($\Delta S_{IC} = \delta AD - TT$) and for DCs ($\Delta S_{DC} = \Delta S_{global} - \Delta S_{IC}$).

Figure 8. Surplus for ICs, DCs and the world at various payments p_c, for $f_r(r) \sim N(-1,1)$, $\delta = 0.5$, $\gamma = 0$, $n = 100$ and $\widehat{BL} = 84$.



As discussed above, the global planner's optimal policy maximizes global efficiency. Hence, the optimal policy from a global planner's perspective would be to set $p_c = \delta = 0.5$. This increases global efficiency by 2.4744 (per 100 plots). However, ICs would not choose to participate in this program. ICs are worse off by 0.0790 (while DCs gain 2.5534). We must restrict ourselves to policies for which $\delta AD - TT \geq 0$. This means that p_c must be reduced below δ. The surplus-maximizing choice for ICs is $p_c = 0.25$. DCs also gain (as long as $p_c > 0$). The policy that gets implemented depends on the relative bargaining power between DCs and ICs.

[22] In fact, p_c in the simulations varies from 0.05 to 1, in steps of 0.05. $\widehat{BL}$ varies from 78 to 93, in steps of 1. A finer grid level would result in a large computation time.

Since ΔS_{DC} is increasing in p_c, any policy with p_c between 0.25 and 0.48 (the price at which ΔS_{IC} = 0) is possible.[23] We rule out policies with p_c < 0.25 since they are Pareto dominated.

This could create a theoretical underpinning for the use of discounting as proposed by policy makers. But there is a second policy lever: Figure 9 shows the impact of changing the baseline generosity on global, IC and DC surplus (fixing p_c = 0.5).

Figure 9. Surplus for ICs, DCs and the world at various assigned baselines $\widehat{BL}$, for $f_r(r)$ ~ N(-1,1), $p_c = \delta = 0.5$, $\gamma = 0$ and $n = 100$.

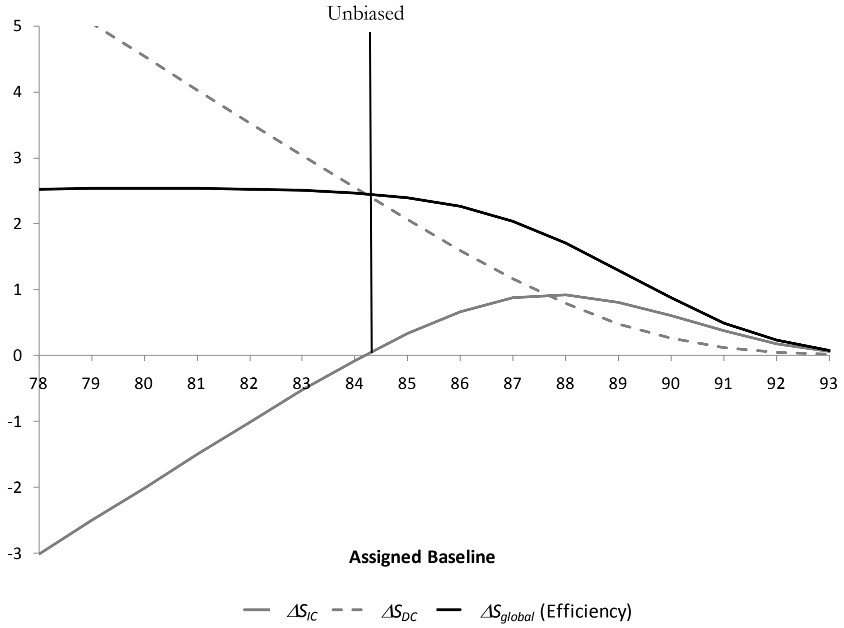


Figure 9 shows that while the global planner's optimal policy would set the baseline as generously as possible, ICs will not choose to participate. In fact, the set of policies that meet the participation constraints and are Pareto optimal consists of $\widehat{BL}$ between 84.25 (the baseline at which ΔS_{IC} = 0) and 88. Again, the final outcome depends on bargaining power.

Figures 8 and 9 demonstrate that if $\gamma = 0$ any policy that ICs will choose to participate in must either involve credit discounting, or setting more stringent than unbiased baselines, or both. It is not straightforward which option should be preferred. On one extreme, policies that leave the ICs indifferent about participating (p_c = 0.48 or $\widehat{BL}$ = 84.25) yield similar surplus for

[23] When the project scale is constrained to n = 10, the planner's efficiency maximizing choice would be $p_c = 0.6 > \delta$. The feasible individually rational set of policies restricts p_c to between 0.10 and 0.26.

DCs. On the other extreme, ICs can extract more surplus by setting p_c = 0.25 than by setting $\widehat{BL}$ = 88.[24]

A third potential policy combines changing baselines and discounting credits. Figure 10 shows contour plots of the surplus for ICs, DCs and the world, when p_c varies between 0 and 0.65, and $\widehat{BL}$ between 82 and 92.

Figure 10. Contour plots of the surplus for ICs, DCs and the world at various baselines $\widehat{BL}$, for $\gamma = 0$, $f_r(r) \sim N(-1,1)$, $n = 100$, at various combinations of (p_c, $\widehat{BL}$).

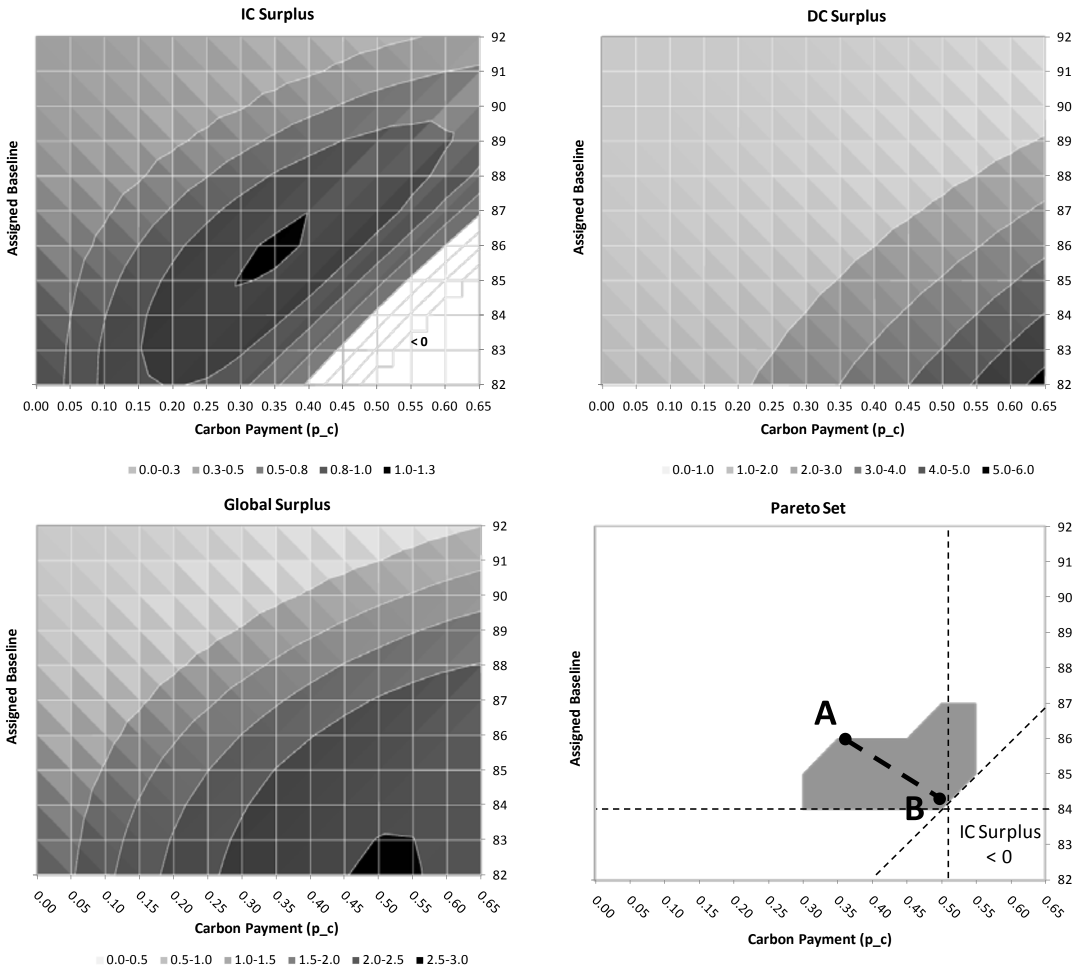


Figure 10 indicates that global surplus is decreasing in $\widehat{BL}$ and increasing when p_c tends towards δ (lower left panel). DCs' surplus is decreasing in $\widehat{BL}$ and increasing in p_c (upper right panel). However, the ICs' optimal choice would be (p_c, $\widehat{BL}$) = (0.35, 86) (upper left panel): a *combination* of credit discounting and a more stringent than unbiased baseline would be optimal from the ICs' perspective. The intuition is that a slight change from $p_c = \delta$ to $p_c = \delta - \varepsilon$ infra-

[24] This could suggest that DCs would be well advised to negotiate on the basis of less generous baselines rather than accepting discounted prices.

marginally reduces all payments for those who opt in, while the change in the opt in probability and the efficiency loss from deforesting plots with returns between δ - ε and δ is small. Similarly, a slight increase in the assigned baseline leads to first order infra-marginal gains but only second-order losses. Hence, we conclude that there is a potential economic rationale for the simultaneous use of discounting and being strict with baselines.

As discussed above, which policy gets implemented ultimately depends on the relative bargaining positions of ICs and DCs. We imposed two constraints: (1) no party can lose relative to no policy, and (2) a policy cannot be Pareto dominated. The lower right panel shows this "conditional Pareto set". The vertical and horizontal dotted lines indicate the environmental externality (δ = 0.5) and the unbiased baseline (BL = 84). The upward-sloping diagonal dotted line marks the area below which IC surplus is less than zero. This figure indicates that, first of all, the global social planner's optimal policy (p_c, $\widehat{BL}$) = (0.50, 0) is not in this feasible set. In addition, the "unbiased policy" (p_c, $\widehat{BL}$) = (0.50, 84) is also not feasible. These policies have a negative surplus for ICs. The ICs' optimal policy (p_c, $\widehat{BL}$) = (0.35, 86) would be obtained if the ICs have all the bargaining power. The more bargaining power lies with DCs, the more the policy will move towards the lower right side of the Pareto set. Point A gives most surplus to ICs, while point B gives most surplus to DCs (while ICs have zero surplus). The dotted line joining point A and B gives some indication of the true Pareto efficient set.[25]

The results discussed above assumed that γ = 0: the net effect of deadweight loss of raising revenues in ICs cancels out against any utility that ICs may receive from transferring funds to poor countries. Changing γ will change the location of the ICs' optimal policy choice. If γ > 0, transfers are more costly for ICs. This moves the feasible policy set further away from the globally efficient solution. For example, when γ = 0.25, the ICs optimal policy choice would be (p_c, $\widehat{BL}$) = (0.25, 85), while, within the Pareto optimal set, DCs are best off with (p_c, $\widehat{BL}$) = (0.35, 84). Although the ICs would optimally choose a baseline that is slightly more generous than if γ = 0, they choose to discount the price much more heavily, resulting in a larger global efficiency loss from the global planner's perspective.[26] The IC surplus from the "unbiased policy" (p_c, $\widehat{BL}$) = (0.50, 84) is now much more negative than if γ = 0. In extreme cases, the Pareto set covers a

[25] The fact that the bottom right quadrant of Figure 10 reports a 2-dimensional Pareto *region* is a result of the discrete grid. As the grid becomes finer, the Pareto set will converge to a curve (approximately between points A and B).
[26] More detailed calculations available from the authors on request.

region with very low p_c and very high $\widehat{BL}$. In that case, no globally meaningful avoided deforestation policy is possible.

In contrast, if $\gamma < 0$, the feasible policy set will move closer to the globally efficient policy. For example, when γ = -0.25, the ICs optimal policy choice would be (p_c, $\widehat{BL}$) = (0.50, 87), while DCs are best off with (p_c, $\widehat{BL}$) = (0.60, 83). The feasible and Pareto efficient policy set is more efficient from the global planner's perspective. The "unbiased policy" (p_c, $\widehat{BL}$) = (0.50, 84) is now feasible. Finally, note that some feasible policies do not require a combination of discounting and stringent baselines – in fact, they do the opposite.

The results in this section have highlighted that under reasonable assumptions ($\gamma \geq 0$), ICs should find it in their best interest to propose a deforestation policy with a combination of discounted credits and baselines that are more stringent than the expected "business as usual" deforestation rate. Moreover, the framework has made explicit the role of benefits and costs of transfers. Depending on negotiators' perception of their country's inherent willingness to transfer (γ), this approach can help them design a "most desired policy". Greater willingness to transfer will generate larger global efficiency gains. The framework also makes explicit the feasible set of policies, which could help DCs determine how far they can bargain with ICs about the division of rents and to what extent they should offer concessions on price relative to baselines.

5. Conclusion and Policy Implications

This paper built a model of landowner decisions and a voluntary avoided deforestation program. It demonstrated the tradeoff between three policy criteria: efficiency, avoided deforestation and payments per unit of avoided deforestation (transfers) when there is asymmetric information. It analyzed the effects of increasing the scale of required projects including when returns are spatially correlated. It then explored the effects on the three policy criteria of two other policy levers: changing the payment per hectare and the baseline. Finally, the paper identifies policies that are individually rational for both industrialized and developing countries and are on the Pareto frontier.

Summary of the main findings:

- We conclude that under almost all circumstances, voluntary deforestation programs (or, in fact, general offset programs) will perform better with increased required scale of project. Only under very special conditions on the return distribution will increasing the required scale of the projects in the program not lead to improved policy outcomes.

- A global social planner would maximize efficiency, but this is not individually rational for ICs.
- If the ICs are averse to transferring funds to developing countries ($\gamma \geq 0$), their preferred contract involves a combination of credit discounting (price below δ) and setting baselines stricter than business-as-usual ($r^* > 0$). These contracts are inefficient. The feasible and Pareto efficient contract set between ICs and DCs contains contracts that involve a combination of credit discounting and stringent baselines.
- As the aversion to transfers γ falls, the tradeoff between efficiency and ICs' willingness to participate diminishes.

Our key messages for policy makers are three. First, make 'projects' as large as possible. Regional or national scale REDD programs where funds or credits are transferred to the government on the basis of aggregated regional or national monitoring data will be much more efficient and offer better value for money. Satellite monitoring of forests is feasible and takes away any ability to cheat. DC governments will not (necessarily) achieve avoided deforestation through smaller-scale voluntary payment programs, but could use a variety of policy instruments to achieve reductions (e.g., strengthening property rights for landowners, abolishing agricultural subsidies).

Second, invest in research to improve understanding of deforestation drivers. This will allow more accurate assessment of returns and hence business as usual deforestation rates. Moreover, this will help identify domestic policies to effectively control deforestation.

Third, the use of credit discounting or below market prices and the use of baselines more stringent than business as usual can be necessary when avoided deforestation policies depend on voluntary contributions from industrialized countries. The level of discounting and stringency required and their optimal mix depends on the developing country marginal cost of abatement (returns distribution), the level of observation errors and the industrialized countries' generosity toward developing countries.

If industrialized countries can be encouraged to be more generous it will be easier to create an efficient international REDD framework. Combined with effective domestic policies that respond to the international incentives, this could meet the expectations of those who promote avoided deforestation as a key climate mitigation option in the short term.

6. References

Andam, K., P. Ferraro, A. Pfaff, J. Robalino, and A. Sanchez. 2008. "Measuring the Effectiveness of Protected-Area Networks in Reducing Deforestation." *Proceedings of the National Academy of Sciences* 105(42): 16089–16094.

Arguedas, C. and D.P. Van Soest. 2009. "On reducing the windfall profits in environmental subsidy programs." *Journal of Environmental Economics and Management* 58: 192 – 205.

Blackman, A., A. Pfaff, J. Robalino, and Y. Zepeda. 2009. Mexico's Natural Protected Areas: Enhancing Effectiveness and Equity. Interim Narrative Report to the Tinker Foundation. Resources for the Future: Washington, DC.

Busch J., B. Strassburg, A. Cattaneo, R. Lubowski, A. Bruner, R. Rice, A. Creed, R. Ashton and F. Boltz. 2009. "Comparing climate and cost impacts of reference levels for reducing emissions from deforestation." *Environmental Research Letters* 4.

Chomitz, K. 2007. "At Loggerheads? Agricultural Expansion, Poverty Reduction, and Environment in the Tropics." World Bank: Washington, DC.

Fischer, C. 2005. "Project-based mechanisms for emissions reductions: balancing trade-offs with baselines." *Energy Policy* 33: 1807 – 1823.

He, G. and R. Morse. 2010. "Making Carbon Offsets Work in the Developing World: Lessons from the Chinese Wind Controversy." Stanford University, Program on Energy and Sustainable Development Working Paper #90.

Kerr, S. 1995. "Adverse Selection and Participation in International Environmental Agreements" in *Contracts and Tradeable Permit Markets in International and Domestic Environmental Protection*, PhD Thesis, Harvard University.

Kerr, S., A.S.P. Pfaff and A. Sanchez. 2002. "The dynamics of deforestation: evidence from Costa Rica." Motu Manuscript, www.motu.org.nz

Kerr, S., J. Hendy, S. Liu and A.S.P. Pfaff. 2004. "Tropical forest protection, uncertainty and the environmental integrity of carbon mitigation policies." Motu Working paper 04-03.

Kindermann, G., M. Obersteiner, B. Sohngen, J. Sathaye, K. Andrasko, E. Rametsteiner, B. Schlamadinger, S. Wunder and R. Beach. 2008. "Global Cost Estimates of Reducing Carbon Emissions Through Avoided Deforestation." *Proceedings of the National Academy of Sciences* 105(30): 10302 – 10307.

Kollmuss, A., M. Lazarus and G. Smith. 2010. "Discounting Offsets: Issues and Options." Stockholm Environment Institute Working Paper WP-US-1005.

Lewis, D.J., A.J. Plantinga, E. Nelson and S. Polasky. 2009. "The efficiency of voluntary incentive policies for preventing biodiversity loss," unpublished manuscript.

Mason, C.F. and A.J. Plantinga. 2010. "The Additionality Problem with Offsets: Optimal Contracts for Carbon Sequestration in Forests." Working paper. Available at http://idei.fr/doc/conf/ere/papers_2010/paper_Mason.pdf.

Melillo, J.M., J.M. Reilly, D.W. Kicklighter, A.C. Gurgel, T.W. Cronin, S. Paltsev, B.S. Felzer, X. Wang, A.P. Sokolov, and C.A. Schlosser. 2009. "Indirect Emissions from Biofuels: How Important?" *Science* 326: 1397 – 1399.

Millard-Ball, A. 2010. "Adverse Selection in an Opt-In Emissions Trading Program: The case of sectoral no-lose targets for transportation." Working paper, Stanford University.

Montero, J.P. 1999. "Voluntary Compliance with Market-Based Environmental Policy: Evidence from the U. S. Acid Rain Program." *Journal of Political Economy* 107(5): 998 – 1033.

Montero, J.P. 2000. "Optimal design of a phase-in emissions trading program." *Journal of Public Economics* 75: 273 – 291.

Pfaff, A., S. Kerr, L. Lipper, R. Cavatassi, B. Davis, J. Hendy, G.A. Sanchez-Azofeifa (2007). "Will buying tropical forest carbon benefit the poor? Evidence from Costa Rica." *Land Use Policy* 24: 600 – 610.

Plantinga, A.J. and K.R. Richards. 2008. "International Forest Carbon Sequestration in a Post-Kyoto Agreement." Discussion Paper 08-11, Harvard Project on International Climate Agreements, Belfer Center for Science and International Affairs, Harvard Kennedy School.

Robalino, J., A. Pfaff, G. A. Sánchez-Azofeifa, F. Alpízar, C. León, and C. M. Rodríguez. 2008. "Deforestation Impacts of Environmental Services Payments: Costa Rica's PSA Program 2000–2005." Environment for Development Discussion Paper 08-24. Washington, DC: Resources for the Future.

Richards, K. and K. Andersson. 2000. "The Leaky Sink: Persistent Obstacles to a Forest Carbon Sequestration Program Based on Individual Projects," *Climate Policy* 1: 41–54.

Sanchez-Azofeifa, A., A. Pfaff, J.A. Robalino and J.P. Boomhower. 2007. "Costa Rica's Payment for Environmental Services Program: Intention, Implementation, and Impact." *Conservation Biology* 21(5): 1165 – 1173.

Stern, N. 2006. *The economics of climate change: The Stern Review*. Cambridge University Press.

Strand, J. 1997. "Developing-country resource extraction with asymmetric information and sovereign debt: a theoretical analysis." *Environment and Development Economics* 2: 265 – 289.

Wise, M., K. Calvin, A. Thomson, L. Clarke, B. Bond-Lamberty, R. Sands, S.J. Smith, A. Janetos, J. Edmonds. 2009. "Implications of Limiting CO2 Concentrations for Land Use and Energy" Science 324: 1183 – 1186.

7. Appendices

Appendix 1 – Proof of Proposition 2

We calculate the efficiency gain $\Delta S_{n=2}$ from 2-plot REDD relative to no policy, and compare this to the efficiency gain $2*\Delta S$ with ΔS defined as $\int_0^{p_c}(p_c - r)\left(\int_{-\infty}^{-r} f_\varepsilon(\varepsilon)d\varepsilon\right)f_r(r)dr$. Assuming symmetric i.i.d. distributions f_r and f_ε, $\Delta S_{n=2}$ equals

$$\begin{aligned}\Delta S_{n=2} = &\int_0^{p_c}\int_0^{p_c}(2\delta - r_1 - r_2)\,Pr\big(\widehat{BL}_2 = 0\big)f_r(r_1)f_r(r_2)dr_2dr_1\\ &+2*\int_0^{p_c}\int_{-\infty}^{0}(\delta - r_1)\,Pr\big(\widehat{BL}_2 \in \{0,1\}\big)f_r(r_1)f_r(r_2)dr_2dr_1\\ &+2*\int_0^{p_c}\int_{p_c}^{\infty}(\delta - r_1)\,Pr\big(\widehat{BL}_2 = 0\big)f_r(r_1)f_r(r_2)dr_2dr_1\\ &+\int_0^{p_c}\int_0^{p_c-r_1}(\delta - r_1)\,Pr\big(\widehat{BL}_2 = 1\big)f_r(r_1)f_r(r_2)dr_2dr_1\end{aligned} \tag{20}$$

These terms refer to the areas A-F in Figure 6. The first and last terms represent areas C + D, the second term represents A + B and the third term represents area E + F. Using

$$\begin{aligned}Pr\big(\widehat{BL}_2 = 0\big) &= \int_{-r_1}^{\infty}\int_{-r_2}^{\infty} f_\varepsilon(\varepsilon_1)f_\varepsilon(\varepsilon_2)d\varepsilon_1\,d\varepsilon_2 = F_\varepsilon(r_1)F_\varepsilon(r_2)\\ Pr\big(\widehat{BL}_2 = 1\big) &= F_\varepsilon(r_1)\big(1 - F_\varepsilon(r_2)\big) + F_\varepsilon(r_2)\big(1 - F_\varepsilon(r_1)\big)\\ Pr\big(\widehat{BL}_2 \in \{0,1\}\big) &= Pr(BL\wedge = 0) + Pr(BL\wedge = 1)\end{aligned} \tag{21}$$

we can rewrite (20) as

$$\begin{aligned}\Delta S_{n=2} = &\,2*\Delta S*\int_0^{p_c}\left(\int_{-r_2}^{\infty} f_\varepsilon(\varepsilon_2)d\varepsilon\right)f_r(r_2)dr_2\\ &+2*\Delta S*\int_{-\infty}^{0}\left(\int_{-r_2}^{\infty} f_\varepsilon(\varepsilon_2)d\varepsilon\right)f_r(r_2)dr_2\\ &+2*\int_0^{p_c}\int_{-\infty}^{0}(\delta - r_1)\,Pr\big(\widehat{BL}_2 = 1\big)f_r(r_1)f_r(r_2)dr_2dr_1\\ &+2*\Delta S*\int_{p_c}^{\infty}\left(\int_{-r_2}^{\infty} f_\varepsilon(\varepsilon_2)d\varepsilon\right)f_r(r_2)dr_2\\ &+\int_0^{p_c}\int_0^{p_c-r_1}(\delta - r_1)\,Pr\big(\widehat{BL}_2 = 1\big)f_r(r_1)f_r(r_2)dr_2dr_1\end{aligned} \tag{22}$$

Which can in turn be simplified to

$$\begin{aligned}
&\Delta S_{n=2} = \\
&+2 * \Delta S * \int_{-\infty}^{\infty} \left(\int_{-r_2}^{\infty} f_\varepsilon(\varepsilon_2) d\varepsilon \right) f_r(r_2) dr_2 \\
&+2 * \int_0^{p_c} \int_{-\infty}^{0} (\delta - r_1)\, Pr\big(\widehat{BL}_2 = 1\big) f_r(r_1) f_r(r_2) dr_2 dr_1 \\
&+ \int_0^{p_c} \int_0^{p_c - r_1} (\delta - r_1)\, Pr\big(\widehat{BL}_2 = 1\big) f_r(r_1) f_r(r_2) dr_2 dr_1
\end{aligned} \tag{23}$$

This can be further simplified to

$$\begin{aligned}
&\Delta S_{n=2} = \Delta S \\
&+2 * \int_0^{p_c} \int_{-\infty}^{0} (\delta - r_1)\, Pr\big(\widehat{BL}_2 = 1\big) f_r(r_1) f_r(r_2) dr_2 dr_1 \\
&+ \int_0^{p_c} \int_0^{p_c - r_1} (\delta - r_1)\, Pr\big(\widehat{BL}_2 = 1\big) f_r(r_1) f_r(r_2) dr_2 dr_1
\end{aligned} \tag{24}$$

Hence, a 2-plot program is more efficient than a single-plot program if the last two terms exceed ΔS. First, let's expand the second term using (20)

$$\begin{aligned}
&2 * \int_0^{p_c} \int_{-\infty}^{0} (\delta - r_1)\, Pr\big(\widehat{BL}_2 = 1\big) f_r(r_1) f_r(r_2) dr_2 dr_1 \\
&= 2 * \int_0^{p_c} \int_{-\infty}^{0} (\delta - r_1) \Big(F_\varepsilon(r_1)\big(1 - F_\varepsilon(r_2)\big) + F_\varepsilon(r_2)\big(1 - F_\varepsilon(r_1)\big) \Big) f_r(r_1) f_r(r_2) dr_2 dr_1 \\
&= 2 * \int_0^{p_c} \int_{-\infty}^{0} (\delta - r_1)\, F_\varepsilon(r_1)\big(1 - F_\varepsilon(r_2)\big) f_r(r_1) f_r(r_2) dr_2 dr_1 \\
&+2 * \int_0^{p_c} \int_{-\infty}^{0} (\delta - r_1)\, F_\varepsilon(r_2)\big(1 - F_\varepsilon(r_1)\big) f_r(r_1) f_r(r_2) dr_2 dr_1 \\
&= 2 * \int_0^{p_c} (\delta - r_1) F_\varepsilon(r_1) \left(\int_{-\infty}^{0} \big(1 - F_\varepsilon(r_2)\big) f_r(r_2) dr_2 \right) f_r(r_1) dr_1 \\
&+2 * \int_0^{p_c} (\delta - r_1)\big(1 - F_\varepsilon(r_1)\big) \left(\int_{-\infty}^{0} F_\varepsilon(r_2)\, f_r(r_2) dr_2 \right) f_r(r_1) dr_1 \\
&> 2 * \frac{3}{8} \int_0^{p_c} (\delta - r_1) F_\varepsilon(r_1) f_r(r_1) dr_1 = \frac{3}{4} \Delta S
\end{aligned} \tag{25}$$

The inequality holds because $\int_{-\infty}^{0} \big(1 - F_\varepsilon(r_2)\big) f_r(r_2) dr_2 = \frac{3}{8}$.

Then, let's expand the third term in (23).

$$
\begin{aligned}
&\int_0^{p_c}\int_0^{p_c-r_1}(\delta-r_1)\,Pr\left(\widehat{BL}_2=1\right)f_r(r_1)f_r(r_2)dr_2dr_1 \\
&=\int_0^{p_c}\int_0^{p_c-r_1}(\delta-r_1)\left(F_\varepsilon(r_1)\left(1-F_\varepsilon(r_2)\right)+F_\varepsilon(r_2)\left(1-F_\varepsilon(r_1)\right)\right)f_r(r_1)f_r(r_2)dr_2dr_1 \\
&=\int_0^{p_c}\int_0^{p_c-r_1}(\delta-r_1)\,F_\varepsilon(r_1)\left(1-F_\varepsilon(r_2)\right)f_r(r_1)f_r(r_2)dr_2dr_1 \\
&\quad+\int_0^{p_c}\int_0^{p_c-r_1}(\delta-r_1)\,F_\varepsilon(r_2)\left(1-F_\varepsilon(r_1)\right)f_r(r_1)f_r(r_2)dr_2dr_1 \\
&=\int_0^{p_c}(\delta-r_1)F_\varepsilon(r_1)\left(\int_0^{p_c-r_1}\left(1-F_\varepsilon(r_2)\right)f_r(r_2)dr_2\right)f_r(r_1)dr_1 \\
&\quad+\int_0^{p_c}(\delta-r_1)\left(1-F_\varepsilon(r_1)\right)\left(\int_0^{p_c-r_1}F_\varepsilon(r_2)\,f_r(r_2)dr_2\right)f_r(r_1)dr_1
\end{aligned}
\tag{26}
$$

Hence, a 2-plot program is more efficient than a single-plot program if

$$
\begin{aligned}
2*\int_0^{p_c}(\delta-r_1)\left(1-F_\varepsilon(r_1)\right)\left(\int_{-\infty}^{0}F_\varepsilon(r_2)\,f_r(r_2)dr_2\right)f_r(r_1)dr_1& \\
+\ \mathit{third\ term}& \\
>\frac{1}{4}\Delta S&
\end{aligned}
\tag{27}
$$

Appendix 2

Table A1. The impact of increasing the required project scale: alternative returns distributions.

$f_r(r)$	1-plot	2-plot	10-plot	100-plot	$\sigma_\varepsilon = 0$
Δefficiency					
N(-1,1)	1.3331	1.4157	1.8248	2.4744	2.5319
N(0,1)	3.0515	3.2681	4.0739	4.8713	4.8854
BMN(0.5,0.1)	1.5379	1.4492	1.2537	1.6858	1.9975
AD					
N(-1,1)	5.3387	5.5368	6.5951	8.9545	9.1848
N(0,1)	13.0382	13.4370	15.6801	19.0864	19.1462
BMN(0.5,0.1)	19.9492	18.7200	15.6702	21.0403	25.0099
TT					
N(-1,1)	4.8817	4.8158	4.5139	4.5562	4.5924
N(0,1)	10.2271	9.9261	9.1712	9.5569	9.5731
BMN(0.5,0.1)	14.0357	13.1049	10.1423	10.9875	12.5058
PAD					
N(-1,1)	0.9149	0.8702	0.6845	0.5089	0.5000
N(0,1)	0.7845	0.7388	0.5849	0.5007	0.5000
BMN(0.5,0.1)	0.7045	0.7001	0.6472	0.5222	0.5000
Opt-in					
N(-1,1)	4.88%	18.30%	58.42%	97.47%	100%
N(0,1)	10.23%	35.53%	78.36%	99.72%	100%
BMN(0.5,0.1)	14.05%	44.53%	65.95%	84.71%	100%

Notes: Each row is based on 10,000 random draws from the probability distributions. $\delta = p_c = 0.5$. $f_\varepsilon(\varepsilon) \sim N(0,0.5)$. *Δefficiency*, *AD* and *TT* are all normalized per 100 plots. Assigned baselines are unbiased.

The effect of scale is the same for the N(0,1) distribution and the central case distribution. Both efficiency, *AD* and *PAD* improve with scale. The BMN(0.5,0.1) distribution demonstrates the results in Section 3.1.2: there exist returns distributions for which increasing scale does not unambiguously improve efficiency. The intuition is that there are many plots with returns around $p_c = \delta$ (as well as returns close to $-p_c$). Therefore, the BMN(0,5,0.1) distribution has many realizations in area C in Figure 6, which can reduce efficiency by Proposition 2 (in the

2-plot model). In fact, we find that a single-plot policy is more efficient than a 2-plot policy. However, such distributions are highly stylized and unlikely to represent true returns distributions.

Figure A1 summarizes this finding by plotting the efficiency gain as a fraction of the efficient ($\sigma_\varepsilon = 0$) solution.

Figure A1. The impact on efficiency of increasing the project scale, for *n* = 1, 2, 10 and 100.

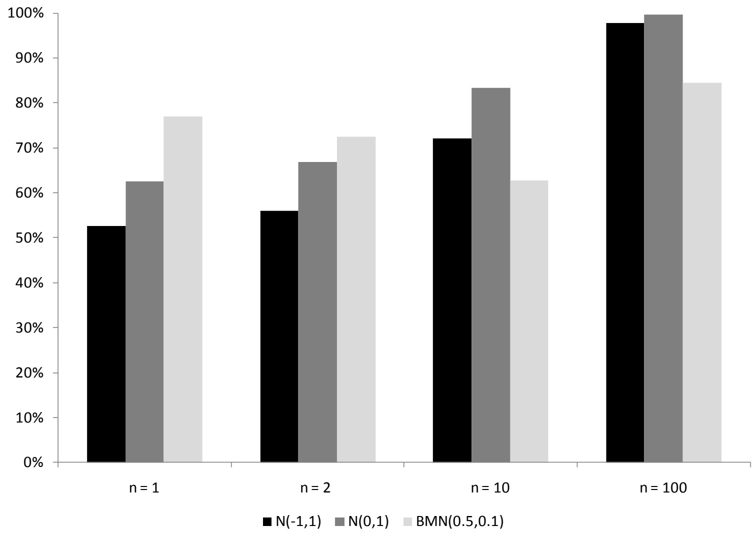


Table A2 shows that the conclusion in Section 3.1.3 (Table 2) that larger required project scale mitigates the adverse impact of spatial correlation also holds for an alternative N(0,1) returns distribution.

Table A2. The impact of spatially correlated returns and observation errors, for n = 1 and n = 100 for an alternative returns distribution.

$f_r(r)$	1-plot	100-	1-plot	100-	1-plot	100-	1-plot	100-	$\sigma_\varepsilon = 0$
	$\rho_r = 0, \rho_\varepsilon = 0$		$\rho_r = 0.5, \rho_\varepsilon = 0.5$		$\rho_r = 0.9, \rho_\varepsilon = 0.9$		$\rho_r = 0.99, \rho_\varepsilon =$		
Δefficienc									
N(-1,1)	1.3331	2.4744	1.3307	2.4130	1.3266	2.0204	1.3267	1.5901	2.5319
N(0,1)	3.0515	4.8713	3.0575	4.8302	3.0517	4.2620	3.0347	3.4793	4.8854
AD									
N(-1,1)	5.3387	8.9545	5.3416	8.7245	5.3317	7.2744	5.3057	5.7233	9.1848
N(0,1)	13.038	19.086	13.055	18.918	13.036	16.618	12.968	13.568	19.146
TT									
N(-1,1)	4.8817	4.5562	4.8950	4.5228	4.8823	4.4151	4.8697	4.5881	4.5924
N(0,1)	10.227	9.5569	10.211	9.5217	10.207	9.2111	10.150	9.5440	9.5731
PAD									
N(-1,1)	0.9149	0.5089	0.9165	0.5184	0.9159	0.6070	0.9185	0.8024	0.5000
N(0,1)	0.7845	0.5007	0.7822	0.5033	0.7830	0.5543	0.7830	0.7037	0.5000
Opt-in									
N(-1,1)	4.88%	97.47%	4.90%	94.96%	4.88%	78.40%	4.87%	37.28%	100%
N(0,1)	10.23%	99.72%	10.21%	98.71%	10.21%	86.78%	10.15%	61.79%	100%

Notes: Each row is based on 10,000 random draws from the probability distributions. $\delta = p_c = 0.5$. σ_ε = 0.5. Assigned baselines are "unbiased".

Motu Working Paper Series

All papers are available online at www.motu.org.nz or by contacting Motu Economic and Public Policy Research.

10-03 Roskruge, Matthew; Grimes, Arthur; McCann, Philip and Poot, Jaques.. 2010. "Social Capital and Regional Social Infrastructure Investment: Evidence from New Zealand".

10-02 Coleman, Andrew. 2010. "Uncovering uncovered interest parity during the classical gold standard era, 1888-1905".

10-01 Coleman, Andrew. 2010. "Squeezed in and squeezed out: the effects of population ageing on the demand for housing".

09-17 Todd, Maribeth and Suzi Kerr. 2009. "How Does Changing Land Cover and Land Use in New Zealand relate to Land Use Capability and Slope?"

09-16 Kerr, Suzi and Wei Zhang. 2009. "Allocation of New Zealand Units within Agriculture in the New Zealand Emissions Trading System".

09-15 Grimes, Arthur; Cleo Ren and Philip Stevens. 2009. "The Need for Speed: Impacts of Internet Connectivity on Firm Productivity".

09-14 Coleman, Andrew and Arthur Grimes. 2009. "Fiscal, Distributional and Efficiency Impacts of Land and Property Taxes".

09-13 Coleman, Andrew. 2009. "The Long Term Effects of Capital Gains Taxes in New Zealand".

09-12 Grimes, Arthur and Chris Young. 2009. "Spatial Effects of 'Mill' Closure: Does Distance Matter?"

09-11 Maré, David C and Steven Stillman. 2009. "The Impact of Immigration on the Labour Market Outcomes of New Zealanders".

09-10 Stillman, Steven and David C Maré. 2009. "The Labour Market Adjustment of Immigrants in New Zealand".

09-09 Kerr, Suzi and Kelly Lock. 2009. "Nutrient Trading in Lake Rotorua: Cost Sharing and Allowance Allocation".

09-08 Coleman, Andrew and Grant M Scobie. 2009. "A Simple Model of Housing Rental and Ownership with Policy Simulations".

09-07 Crawford, Ron. 2009. "Variations in Earnings Growth: Evidence from Earnings Transitions in the NZ Linked Income Survey".

09-06 Maré, David C. 2009. "Agglomeration Elasticities in New Zealand".

09-05 Poot, Jacques. 2009. "Trans-Tasman Migration, Transnationalism and Economic Development in Australasia".

09-04 Hengyun, Ma; Les Oxley and John Gibson. 2009. "China's Energy Situation and Its Implications in the New Millenium".

09-03 Hengyun, Ma; Les Oxley and John Gibson. 2009. "Testing for Energy Market Integration in China".

09-02 Hengyun, Ma; Les Oxley, John Gibson and Bongguen Kim. 2009. "China's Energy Economy: Technical Change, Factor Demand and Interfactor/Interfuel Substitution".

09-01 Cumming, Jackie; Steven Stillman and Michelle Poland. 2009. “Who Pays What for Primary Health Care? Patterns and Determinants of the Fees Paid by Patients in a Mixed Public-Private Financing Model”.

08-14 Coleman, Andrew. 2008. “Tax, Credit Constraints, and the Big Costs of Small Inflation”.

08-13 Coleman, Andrew. 2008. “Storage Under Backwardation: A Direct Test of the Wright-Williams Conjecture”.

08-12 Maré, David C. 2008. “Labour Productivity in Auckland Firms”.

08-11 Stillman, Steven; Malathi Velamuri and Andrew Aitken. 2008. “The Long-Run Impact of New Zealand's Structural Reform on Local Communities”.

08-10 Grimes, Arthur and Andrew Aitken. 2008. “Water, Water Somewhere: The Value of Water in a Drought-Prone Farming Region”.

08-09 Coleman, Andrew. 2008. “Inflation and the Measurement of Saving and Housing Affordability”.

08-08 Coleman, Andrew and Özer Karagedikli. 2008. “The Relative Size of New Zealand Exchange Rate and Interest Rate Responses to News”.

08-07 Grimes, Arthur and Yun Liang. 2008. “Bridge to Somewhere: The Value of Auckland's Northern Motorway Extensions".

08-06 Stillman, Steven and David C Maré. 2008. “Housing Markets and Migration: Evidence from New Zealand”.

08-05 Lock, Kelly and Suzi Kerr. 2008. “Nutrient Trading in Lake Rotorua: Choosing the Scope of a Nutrient Trading System”.

08-04 Kerr, Suzi and Andrew Sweet. 2008. “Inclusion of Agriculture and Forestry in a Domestic Emissions Trading Scheme: New Zealand’s Experience to Date”.

08-03 Kerr, Suzi and Kit Rutherford. 2008. “Nutrient Trading in Lake Rotorua: Reporting and Modelling Net Nutrient Inputs”.

08-02 Lock, Kelly and Suzi Kerr. 2008. “Nutrient Trading in Lake Rotorua: Overview of a Prototype System”.

08-01 Stroombergen, Adolf. 2008. “ESSAM General Equilibrium Model: Estimation of 2005/06 Input-Output Tables”.

07-13 Hall, Viv and John McDermott. 2007. "A quarterly post-World War II real GDP series for New Zealand".

07-12 Gibson, John; Trinh Le and Steven Stillman. 2007. "What Explains the Wealth Gap Between Immigrants and the New Zealand Born?"

07-11 Maré, David C.; Melanie Morten and Steven Stillman, 2007. “Settlement Patterns and the Geographic Mobility of Recent Migrants to New Zealand”.

07-10 Grimes, Arthur; David C. Maré and Melanie Morten, 2007. “Adjustment in Local Labour and Housing Markets.”

07-09 Grimes, Arthur and Yun Liang, “Spatial Determinants of Land Prices in Auckland: Does the Metropolitan Urban Limit Have an Effect?”

07-08 Kerr, Suzi; Kit Rutherford and Kelly Lock, “Nutrient Trading in Lake Rotorua: Goals and Trading Caps”.

07-07 Hendy, Joanna; Suzi Kerr and Troy Baisden, "The Land Use in Rural New Zealand Model Version 1 (LURNZ v1): Model Description".

07-06 Lock, Kelly and Suzi Kerr, "Nutrient Trading in Lake Rotorua: Where Are We Now?"

07-05 Stillman, Steven and David C Maré, "The Impact of Immigration on the Geographic Mobility of New Zealanders".

07-04 Grimes, Arthur and Yun Liang, "An Auckland Land Value Annual Database".

07-03 Kerr, Suzi; Glen Lauder and David Fairman, "Towards Design for a Nutrient Trading Programme to Improve Water Quality in Lake Rotorua".

07-02 Lock, Kelly and Stefan Leslie, "New Zealand's Quota Management System: A History of the First 20 Years".

07-01 Grimes, Arthur and Andrew Aitken. 2007. "House Prices and Rents: Socio-Economic Impacts and Prospects".

06-09 Maani, Sholeh A.; Rhema Vaithianathan and Barbara Wolf. 2006. "Inequality and Health: Is House Crowding the Link?"

06-08 Maré, David C and Jason Timmins. 2006. "Geographic Concentration and Firm Productivity".

06-07 Grimes, Arthur; David C Maré and Melanie Morten. 2006. "Defining Areas Linking Geographic Data in New Zealand".

06-06 Maré, David C and Yun Liang. 2006. "Labour Market Outcomes for Young Graduates".

06-05 Hendy, Joanna and Suzi Kerr. 2006. "Land-Use Intensity Module: Land Use in Rural New Zealand Version 1".

06-04 Hendy, Joanna; Suzi Kerr and Troy Baisden. 2006. "Greenhouse Gas Emissions Charges and Credits on Agricultural Land: What Can a Model Tell Us?"

06-03 Hall, Viv B; John McDermott and James Tremewan. 2006. "The Ups and Downs of New Zealand House Prices".

06-02 McKenzie, David; John Gibson and Steven Stillman. 2006. "How Important is Selection? Experimental vs Non-Experimental Measures of the Income Gains from Migration".

06-01 Grimes, Arthur and Andrew Aitken. 2006. "Housing Supply and Price Adjustment".

05-14 Timmins, Jason. 2005. "Is Infrastructure Productive? Evaluating the Effects of Specific Infrastructure Projects on Firm Productivity within New Zealand".

05-13 Coleman, Andrew; Sylvia Dixon and David C Maré, "Māori Economic Development—Glimpses from Statistical Sources".

05-12 Maré, David C 2005. "Concentration, Specialisation and Agglomeration of Firms in New Zealand".

05-11 Holmes, Mark J and Arthur Grimes. 2005. "Is There Long-Run Convergence of Regional House Prices in the UK?"

05-10 Hendy, Joanna and Suzi Kerr. 2005. "Greenhouse Gas Emission Factor Module: Land Use in Rural New Zealand—Climate Version 1".

05-09 Poland, Michelle and David C Maré. 2005. "Defining Geographic Communities".

05-08 Kerr, Suzi; Joanna Hendy, Emma Brunton and Isabelle Sin. 2005. "The Likely Regional Impacts of an Agricultural Emissions Policy in New Zealand: Preliminary Analysis".

05-07 Stillman, Steven. 2005. "Examining Changes in the Value of Rural Land in New Zealand between 1989 and 2003".

05-06 Dixon, Sylvia and David C Maré. 2005. "Changes in the Māori Income Distribution: Evidence from the Population Census".

05-05 Sin, Isabelle and Steven Stillman. 2005. "The Geographical Mobility of Māori in New Zealand".

05-04 Grimes, Arthur. 2005. "Regional and Industry Cycles in Australasia: Implications for a Common Currency".

05-03 Grimes, Arthur. 2005. "Intra and Inter-Regional Industry Shocks: A New Metric with an Application to Australasian Currency Union".

05-02 Grimes, Arthur; Robert Sourell and Andrew Aitken. 2005. "Regional Variation in Rental Costs for Larger Households".

05-01 Maré, David C. 2005. "Indirect Effects of Active Labour Market Policies".

04-12 Dixon, Sylvia and David C Maré. 2004. "Understanding Changes in Māori Incomes and Income Inequality 1997–2003".

04-11 Grimes, Arthur. 2004. "New Zealand: A Typical Australasian Economy?"

04-10 Hall, Viv and John McDermott. 2004. "Regional Business Cycles in New Zealand: Do They Exist? What Might Drive Them?"

04-09 Grimes, Arthur; Suzi Kerr and Andrew Aitken. 2004. "Bi-Directional Impacts of Economic, Social and Environmental Changes and the New Zealand Housing Market".

04-08 Grimes, Arthur and Andrew Aitken. 2004. "What's the Beef with House Prices? Economic Shocks and Local Housing Markets".

04-07 McMillan, John. 2004. "Quantifying Creative Destruction: Entrepreneurship and Productivity in New Zealand".

04-06 Maré, David C and Isabelle Sin. 2004. "Māori Incomes: Investigating Differences Between Iwi".

04-05 Kerr, Suzi; Emma Brunton and Ralph Chapman. 2004. "Policy to Encourage Carbon Sequestration in Plantation Forests".

04-04 Maré, David C. 2004. "What do Endogenous Growth Models Contribute?"

04-03 Kerr, Suzi; Joanna Hendy, Shuguang Liu and Alexander S. P. Pfaff. 2004. "Uncertainty and Carbon Policy Integrity".

04-02 Grimes, Arthur; Andrew Aitken and Suzi Kerr. 2004. "House Price Efficiency: Expectations, Sales, Symmetry".

04-01 Kerr, Suzi; Andrew Aitken and Arthur Grimes. 2004. "Land Taxes and Revenue Needs as Communities Grow and Decline: Evidence from New Zealand".

03-19 Maré, David C. 2003. "Ideas for Growth?"

03-18 Fabling, Richard and Arthur Grimes. 2003. "Insolvency and Economic Development: Regional Variation and Adjustment".

03-17 Kerr, Suzi; Susana Cardenas and Joanna Hendy. 2003. "Migration and the Environment in the Galapagos: An Analysis of Economic and Policy Incentives Driving Migration, Potential Impacts from Migration Control, and Potential Policies to Reduce Migration Pressure".

03-16 Hyslop, Dean R and David C Maré. 2003. "Understanding New Zealand's Changing Income Distribution 1983–98: A Semiparametric Analysis".

03-15 Kerr, Suzi. 2003. "Indigenous Forests and Forest Sink Policy in New Zealand".

03-14 Hall, Viv and Angela Huang. 2003. "Would Adopting the US Dollar Have Led to Improved Inflation, Output and Trade Balances for New Zealand in the 1990s?"

03-13 Ballantyne, Suzie; Simon Chapple, David C Maré and Jason Timmins. 2003. "Movement into and out of Child Poverty in New Zealand: Results from the Linked Income Supplement".

03-12 Kerr, Suzi. 2003. "Efficient Contracts for Carbon Credits from Reforestation Projects".

03-11 Lattimore, Ralph. 2003. "Long Run Trends in New Zealand Industry Assistance".

03-10 Grimes, Arthur. 2003. "Economic Growth and the Size & Structure of Government: Implications for New Zealand".

03-09 Grimes, Arthur; Suzi Kerr and Andrew Aitken. 2003. "Housing and Economic Adjustment".

03-07 Maré, David C and Jason Timmins, “Moving to Jobs”.

03-06 Kerr, Suzi; Shuguang Liu, Alexander S. P. Pfaff and R. Flint Hughes. 2003. “Carbon Dynamics and Land-Use Choices: Building a Regional-Scale Multidisciplinary Model”.

03-05 Kerr, Suzi. 2003. “Motu, Excellence in Economic Research and the Challenges of ‘Human Dimensions’ Research”.

03-04 Kerr, Suzi and Catherine Leining. 2003. “Joint Implementation in Climate Change Policy”.

03-03 Gibson, John. 2003. “Do Lower Expected Wage Benefits Explain Ethnic Gaps in Job-Related Training? Evidence from New Zealand”.

03-02 Kerr, Suzi; Richard G Newell and James N Sanchirico. 2003. “Evaluating the New Zealand Individual Transferable Quota Market for Fisheries Management”.

03-01 Kerr, Suzi. 2003. “Allocating Risks in a Domestic Greenhouse Gas Trading System”.