Title
Biochar for sustainable agricultural intensification: technical/economic potential, and technology adoption

Author
Crane-Droesch, Andrew

Publication Date
2015

Peer reviewed|Thesis/dissertation

Biochar for sustainable agricultural intensification: technical/economic potential, and technology adoption

By

Andrew Noel Crane-Droesch

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Energy & Resources

in the

Graduate Division

of the

University of California, Berkeley

Committee in Charge:

Professor David I. Levine (Co-chair)
Professor Daniel M. Kammen (Co-chair)
Professor Margaret Torn
Professor Edward Miguel

Summer 2015

Abstract

Biochar for sustainable agricultural intensification: technical/economic potential, and technology adoption

by
Andrew Noel Crane-Droesch

Doctor of Philosophy in Energy & Resources

University of California, Berkeley

Professor David I. Levine, co-chair

Professor Daniel M. Kammen, co-chair

Growing population, changing climate, and human development will require sustainable agricultural intensification in sub-Saharan Africa – a region where most people still live in rural areas, and rural poverty remains severe. While the 20th century saw massive increases in agricultural productivity over most of the world, sub-Saharan Africa was largely bypassed. In many respects, catching up in the 21st century poses more difficult challenges than were faced in the 20th – with the novel challenges of climate change, soil degradation, and the closing agricultural frontier being chief among them. Solutions are required that profitably improve productivity, strengthen and/or rebuild soil fertility, and do so within the limits imposed by a warming and carbon-constrained world. This will be needed if the region is to shake off the stagnation that has characterized it for the past century, and contribute to solving to the global challenges posed by the coming century.

This dissertation focuses on one potential solution – biochar – and follows it from agronomic efficacy, to preliminary economic analysis, to rigorous trial in the field. As such, it seeks to be an example of the interdisciplinary approach that I argue is needed to guide the development, deployment, and scaling of solutions to problems in the environment/development space.

The first chapter is a meta-analysis of crop yield response to biochar. Using data from 84 studies, I (and co-authors) employ meta-analytical, missing data, and semiparametric statistical methods to explain heterogeneity in crop yield responses across different soils, biochars, and agricultural management factors, and then estimate potential changes in yield across different soil environments globally. We find that soil cation exchange capacity and organic carbon were strong predictors of yield response, with low cation exchange and low carbon associated with positive response. We also find that yield response increases over time since initial application, compared to non-biochar controls. High reported soil clay content and low soil pH were weaker predictors of higher yield response. No biochar parameters in our dataset – biochar pH, percentage carbon content, or temperature of pyrolysis – were significant predictors of yield impacts. Projecting our fitted model onto a global soil database, we find the largest potential increases in areas with highly weathered

soils, such as those characterizing much of the humid tropics. Richer soils characterizing much of the world's important agricultural areas appear to be less likely to benefit from biochar.

The second chapter is a preliminary economic analysis of biochar's potential in two contexts – rural western Kenya, and northern Vietnam. Using recall-based datasets from smallholder farmers, I (and co-authors) estimate yields as a function of biochar and fertilizer use. We find an positive association between biochar use and average yields in Kenya, but no correlation in Vietnam. We then use these estimates to calculate optimal input mixes under hypothetical biochar and carbon prices, given heterogeneity in response both to biochar and fertilizer, and heterogeneous budget constraints. In Kenya, we find that biochar is more-likely-than-not to be profitable to adopt for 23% of our sample if unsubsidized and available at its current sale price of around $188/ton, while a hypothetical carbon subsidy of $100/ton CO_2e increases this proportion to 47%, though these proportions are not different from zero at 95% confidence. Because of limited short-term complementarity between biochar and inorganic fertilizer, we estimate that biochar adoption would change profits little, given budget constraints for agricultural inputs. We conclude that carbon subsidies may have a marginal impact on biochar's profitability in Western Kenya, but that further research is needed to improve the precision of these estimates, extend them to account for any longer-term changes in soil characteristics that might impact biochar's profitability, and account for any potential biases stemming from time-varying variables that not measured or modeled in the context of this study.

The third chapter reports the results of a Kenyan field experiment on adoption and impact of biochar, which was motivated by the encouraging findings of the previous two studies. In addition to technical efficacy, I sought to determine what mix of policies might most effectively speed biochar dissemination, given the slow pace of technological change in African agriculture over the past several decades. I randomly assigned prices, demonstrations, and risk-free trials. Yields increased by 37% and 50% in the first two seasons, and response to inorganic fertilizer improved. However, uptake was 2.6% and 10% respectively. Farmers were highly price-sensitive. Social network effects were marginally significant, but positive at low penetration and negative at high penetration. Given uptake well below the social optimum, subsidies for biochar appear justified from a social cost/benefit standpoint.

The dissertation closes with a short discussion of lessons learned, and ways forward for further applied research.

"The only true wisdom is knowing that you know nothing."

Socrates

"How is education supposed to make me feel smarter? Besides, every time I learn something new, it pushes some old stuff out of my brain. Remember when I took that home winemaking course, and I forgot how to drive?"

Homer Simpson

Acknowledgements

My dissertation owes thanks to my advisors; David Levine, Dan Kammen, Margaret Torn, and Ted Miguel. The Kenya project underlying my dissertation owes thanks to Paul Manda, Salim Shaban, Eric Solomonson, David Guerena, Daniel Agness, Allen Baumgardner-Zuzik, Claudia Casaroto, and many, many others. I'm grateful to fellow students, seminar participants, and others from whom I've gotten feedback, criticism, and support. And I'm especially grateful to those who I've forgotten to mention here.

Contents

Mostly to Amanda.
Also to Muuaji the cat, who sat on my lap as I wrote this thing

Chapter 1

The Necessity of Sustainable Intensification in African Agriculture

Global change casts old problems in a new light. This is particularly true with African agriculture, which, even in today's rapidly urbanizing world, still employs most people on the continent. In the period since independence – more than a half-century – African agriculture has largely stagnated. Productivity remains low, rural poverty remains severe, and the technological advancements in agriculture that have contributed to poverty alleviation in much of the rest of the world have largely passed Africa by. Global change makes agricultural intensification – already a critically important and largely unmet goal – more important than ever. Furthermore, global change renders the past is an imperfect guide to the future. While recent years have seen a welcome re-focusing of attention towards African agriculture by governments, NGOs, and the international development community after decades of neglect, simply righting the wrongs of past policy will not be sufficient to address the challenges (and meet a few opportunities) posed by global change – particularly climate change but also other global environmental change, along with economic and demographic shifts.

This introduction begins by briefly recapitulating the state of agriculture in Sub-Saharan Africa, along with the case for agricultural intensification in the region. I will then examine global trends – economic, demographic, and environmental – and discuss what they mean for both the prospects of and the importance of agricultural development in the region. I'll conclude with an introduction to the rest of my dissertation, which I conceptualize as an elaboration of one potential solution – biochar – which the dissertation follows along the course from scientific proof-of-concept, to preliminary economic analysis, to rigorous trial in the field among smallholder farmers. The dissertation concludes with an integrative chapter, summarizing what we've learned about both biochar in particular, and more generally about means of facilitating the transformation in African agriculture to meet some of the challenges posed by global change.

African Agriculture in 2015

For the past several millennia, agriculture has been the fundamental human activity – it is that upon which our civilization is based. While this is true throughout the world (beyond a few remaining hunter-gatherer societies), in few places is it as apparent as it is throughout sub-Saharan Africa. About 75% of the continent still lives in rural areas, on par with South Asia, but substantially higher than the rest of the world. Furthermore, while most of the world has experienced steady gains in agricultural productivity over this period, Africa has largely been left behind, contributing to persistent rural poverty in the region [1]. Drivers of Africa's agricultural stagnation are manifold, inter-related, and operate on different scales. At the macro and meso levels, dysfunctional institutions, poor infrastructure, and market failures constrain the options available to farmers [2]. Some of the most important constraints however are at the micro level, and include imperfect information, uninsured risk, and iliquidity [3]. In particular, stagnation is often attributed to lower use of inorganic fertilizers and hybrid seed [4], which in turn is intertwined with Africa's post-independence economic history[5].

Natural resource degradation – central to this dissertation – is among the most important constraints that African farmers face, given that it mediates most other policy questions that relate to agricultural intensification in the region. Particularly in the weathered soils that comprise most of the region, response to inorganic fertilizer is commonly strongly conditional on soil organic carbon[6, 7]. This can make fertilizers unprofitable, which in turn can lower demand, which in turn can increase costs to suppliers, which increases prices. In other contexts, fertilizer may be profitable in expectation, but its profitability in a given year may be quite variable[8]. For poor households with inadequate safety nets, such gambles can be unacceptable. This dynamic can lead farmers to take more nutrients from their soils than are returned [9, 10], lowering soil organic carbon over time, lowering investment in nutrients, lowering yields, deepening poverty, etc. Several authors have characterized this as a natural resources poverty trap [11, 12].

Trends

While African agriculture has changed somewhat over the past half-century[2] the world is changing fundamentally. Human population is on course to reach 10 billion by the end of the century, while increased demand for animal products among wealthier people means that food demand is expected to rise even faster[13]. Africa is expected to contribute most to this population growth[14]. Given poor trade links with the rest of the world and likely tighter external demand in the foreseeable future[15], it will become essential for the region to produce more food.

This will take place under a backdrop of dramatic changes to the climate in the coming century. Absent adaptation, current estimates put yield decreases between 7-27% by mid-century [16]. However, maladapatation – defined as responses to climate change that exacerbate the problem – seem frighteningly possible. Increased climate risk may make investments in fertilizer riskier and thus less desirable, while poverty itself may exacerbate risk aversion and disinclination to invest.

Land scarcity also looms in many regions. On average, the proportion land under agriculture in Africa has grown by 6% in the period since independence, and about one third of this has been into forest [17]. Durations of fallows have decreased and cultivation has extended into ever-more marginal land [18]. As such, attempting to increase production via extensification may be both more difficult in the 21st century than it was in the 20th, while the increasing salience of forest preservation given climate change makes it even less desirable.

Agricultural Intensification at Baseline

It is clear that African agricultural output must rise if anti-poverty goals are to be met. Many smallholder farmers are net-buyers of food, and increasing production profitably represents less expenditure on food[19]. There is also substantial evidence that GDP growth in agriculture is more effective in reducing poverty than GDP growth in other sectors[2]. Given that further extensification is either impossible or undesirable in much of the region, production per unit land will need to increase. Yet the discourse surrounding agricultural intensification in Africa has a long history, from colonial misunderstandings of the ecology and the economics of land-abundant farming systems [20, 21], to misguided and heavy-handed colonial (and post-colonial) mega-projects[22], to efforts to replicate Asia's green revolution in the 1960s and 1970s, to more recent efforts aimed at sustainable natural resource management. While older efforts were often misguided, changing conditions since the colonial period have made traditional techniques – such as shifting cultivation – less sustainable and/or feasible.

Global change and Agricultural Intensification in Africa

Global change will make agricultural intensification in Africa more urgent than ever. Converging economics and environmental changes – including both the strengthening of existing dynamics as well as more novel pressures – will render currently sub-optimal pathways moreso. On the other hand, global change offers a few opportunities for stimulating sustainable development, along with some opportunities to contribute towards efforts to address global challenges. This section briefly sketches several of these dynamics.

Climate variability, soil carbon, and water storage

Throughout much of Africa, convective rainfall generates most of the precipitation on which rainfed agriculture relies. These events are commonly quite heavy, and can lead to substantial erosion and runoff. At the same time, a few events can comprise an entire year's rainfall in certain areas. Where there are large gaps between rainfall events within the growing season – or where there is too much rainfall within a short period, crop performance can suffer.

Rainfall events are expected to grow more intense as climate change progresses [23]. This increase in intensity is largely absent from many assessments of the effect of climate change on agriculture that are based on GCM output, as GCMs typically can't resolve fine-scale dynamics inside of convective rainfall systems[24]. This means that many of the assessments of the impact of climate change on agriculture that are based on GCM output may be overly optimistic.

Soil carbon – variously generated on farms from compost, crop residues, agroforestry, or biochar – can improve soil hydraulic properties, particularly water-holding capacity. In wet conditions, it can facilitate infiltration of water in soil, including drainage downwards – thereby reducing runoff and mitigating erosive pressures. In dry conditions, soil carbon can serve as a sponge – retaining water that might otherwise be lost from the system. As rainfall grows more intense, these properties will increase in value.

Soil carbon and fertilizer-use efficiency

In tropical soils, yield response to fertilizer is commonly mediated by soil carbon content. Given relatively high prices for inorganic fertilizer throughout the continent, along with their marginal profitability in some areas[6, 7, 25], soil carbon may be a prerequisite to more-sustainable intensification. As such, augmentation of soil carbon levels can be an effective approach to agricultural intensification at baseline, irrespective of climate change. However, production of inorganic fertilizer is commonly greenhouse-gas intensive, and improving returns to fertilizer can reduce the greenhouse gas intensity of agricultural production.

Certain fertilizer-specific factors may become increasingly relevant over time as well. Production of ammonia-based fertilizer generally relies on natural gas as a raw material. While production of natural gas has increased dramatically in recent years, many developed countries have begun to burn more of it as fuel. The effect of this dynamic on fertilizer prices has been ambiguous, but a tighter market could plausibly lead to more volatility in the future. Given that African farmers (and governments) may be more exposed to downside risk than they are able to seize on upside opportunities, improvement of fertilizer-use efficiency can buffer any price shocks by reducing absolute demand per unit crop produced. Finally, any carbon pricing scheme adopted by fertilizer-producing or gas-producing countries may increase the cost of nitrogenous fertilizers.

While nitrogen is derived from ambient air (through the use of energy), and combined with hydrogen (usually derived from natural gas) to form ammonia, phosphate is a non-renewable, mined resource. Unlike fossil fuels, there is no substitute for phosphorous. And unlike many other mined minerals, recycling is generally infeasible. While mainstream estimates do not suggest that "peak phosphorous" will arrive anytime within the next several decades [26], dwindling phosphate reserves and growing demand imply that prices must increase over the long term, absent major new discoveries or technological innovation in phosphorous recovery. Given that tropical soils tend to be more phosphorous-limited than they are nitrogen-limited, African agriculture will be more exposed to these dynamics over the long term than most regions. More-intensive agricultural systems use phosphorous more efficiently, and improving this intensity will buffer the region against this longer-term trend.

Finally, increasing weather risk implies increasing investment risk in fertilizer. For poor farmers, who currently use little fertilizer in large part because of risk [8] this dynamic of potentially higher prices and certainly higher risk suggests that efforts to increase fertilizer use might face a serious headwind. Augmenting soil carbon may help to mitigate this risk by increasing fertilizer use efficiency – effectively lowering its cost per unit output – while buffering against weather risk.

Intensification and Mitigation

While agricultural intensification in sub-Saharan Africa will almost certainly require use of inorganic fertilizers to deliver nutrients to degraded soils[4, 27], and while these fertilizers will tend to be greenhouse-gas intensive, some aspects of agricultural intensification may actually serve to mitigate climate change. Agroforestry and biochar both divert carbon from the atmosphere into the soil through biomass. While these inputs will eventually re-mineralize, many agricultural techniques aimed at bolstering soil carbon would not reach equilibrium for many years, or even centuries to millennia in the case of biochar[28]. Meanwhile, by improving response to fertilizer, soil carbon decreases carbon emissions per unit agricultural output. Over the very long term – multiple centuries – widespread use of biochar, coupled with lower-GHG-intensity fertilizers and tighter nutrient cycling in agroecosystems could represent a means of drawing down atmospheric concentrations of carbon dioxide.

Africa and the World

On the whole, Africa imports more food than it exports[1]. Africa also has substantial yield gaps – technically, yields could improve with better inputs and management[29]. This presents both a risk and an opportunity as food demand and climate change-related disruptions to food supplies increase in the coming century.

This imbalance is a risk inasmuch as shocks to global food supplies far removed from Africa can affect Africa. The 2008 and 2011 food crises saw prices rise dramatically, leading to unrest in much of the region. Droughts in Australia and Russia as well as floods in South Asia contributed to price rises, as did biofuel policies in the US and EU[30] – which can be considered to be (potentially misguided) responses to climate change. In response, several large exporters (Russia and India, in particular) imposed export bans, exacerbating the crisis elsewhere.

While it has been argued that some countries in the region should de-emphasize agricultural development in favor of non-agricultural export market development[31], these decisions should weigh likely scenarios of increasing climate-driven and demand-driven volatility in international food markets, as they are likely to affect the relative desirability of a focus on agricultural development versus reliance on imports and focus on non-agricultural export sector development.

This imbalance is an opportunity inasmuch as the technical potential exists to increase agricultural production in the region – dramatically in some cases [29]. This is in contrast to many developed countries, where technical potential for yield increases is commonly much lower. While production is indeed threatened by climate change, many climate change adaptation measures are also measures that facilitate agricultural intensification, such as the enhancement of soil carbon, establishment of farm trees, and even the development of rural roads and grain banks, which could help buffer against shocks at the same time as they could enable surpluses to reach world markets. Together, this means that the returns to effectively increasing food production throughout the continent may be higher than they had been previously.

The Way Forward

At the level of national policy, measures to stimulate the intensification of agriculture in response to global change are likely to be political questions as much as they are questions of policy – different actors may have different priorities, and urban populations often have disproportionate political power in the region. At the micro level however, it is not always clear what mix of approaches actors in the agricultural sector should take. While there has been substantial agronomic and agroecological research in the past several decades on intensive agricultural practices suited to the sorts of agroecosystems found in sub-Saharan Africa, this research has too-often been divorced from agricultural economics, which until recently has focused almost entirely on the technologies used in the Asian and Latin American green revolutions – inorganic fertilizer, hybrid seed, and irrigation.

Thus, while there has been substantial research on such interventions as agroforestry, compost, and biochar, it has almost entirely been done by physical or biological scientists, and rarely considered economic or behavioral issues. These matter as much as technical efficacy in stimulating uptake and ultimately transforming agriculture. Many technologies that are consistent with sustainable, intensive agriculture are too labor-intensive, and as such are uneconomical for labor constrained households. Other technologies can be too costly when economies of scale are not met. However, lack of deployment is not proof of inappropriateness. African farmers, like most people, are poorly modeled as classical economic actors – they have imperfect information, bounded rationality, lack access to credit, are unable to internalize the externalities of all of their actions, and tend to be justifiably averse to risk. These factors can and do explain much of the absence of technologies that could be beneficial, and addressing these factors has an important role to play in stimulating agricultural intensification.

Meeting the demands of the coming century will require interdisciplinary approaches to research and development. The obvious first step for any given technology is scientific proof-of-concept. This should be followed by basic economic analysis, which in turn should be followed by rigorous trial in the field. Unfortunately, most technologies developed aimed at sustainable agricultural intensification stop after the first step, while much work in economics – in which randomized controlled trials are increasingly the norm in applied research – too often have fail to interact with the scientific community that has been engaged with these questions. While impact evaluations conducted by agricultural and development economists have proliferated over the years, relatively few of these have explicitly focused on the sorts of technologies needed to achieve sustainable agricultural intensification in the region.

Biochar as a Case

This dissertation focuses on biochar, and follows the framework described above: scientific proof of concept, preliminary economic analysis, and rigorous trial in the field. Biochar has been the subject of substantial agronomic and biogeochemical research over the past decade, owing to its ability to boost crop yields while slowing the decomposition of its feedstock biomass. Particularly in weathered soils, yield increases in response to biochar have been dramatic. This, together with the facts that carbon cycling tends to be faster in the tropics, and that tropical agricultural soils are commonly carbon-constrained, immediately suggests that biochar might play a role in stimulating sustainable agricultural intensification.

The first paper is a meta-analysis of yield response to biochar[32]. A proof-of-concept to test whether available data indeed supports the notion that biochar may be agronomically effective in weathered soils, I take data from several published studies, and estimate average yield response to biochar as a function of various soil and biochar properties. I find that the main predictors of high yield response are low cation exchange capacity and low soil organic carbon content. Projecting the model onto a global dataset of soil properties, I find that biochar is likely to have substantially higher benefit in the global south.

The second paper [33] is a preliminary economic analysis of biochar, based on data collected when I was first scoping my main dissertation project in Bungoma, Kenya. I collected a recall-based panel dataset of 300 early biochar adopters, and use this to estimate yield response as well as complementarity with inorganic fertilizer. I also incorporate a similar dataset from early biochar adopters in several regions of Vietnam, which provides a stark contrast – both economically and agroecologically – with Western Kenya. I find large yield increases in the Kenya sample, but no effect in the Vietnam dataset. Using my estimates of impact, I then estimate the impact of a carbon price on optimal biochar adoption for the Kenya sample. While I find some suggestive evidence that a price on carbon – in the form of a subsidy for farmers – might stimulate biochar adoption for a profit-maximizing farmer, results are uncertain, and the paper calls for further research on biochar in smallholder farming systems, using more rigorous methods in which establishment of causal relationships is more straightforward.

The final paper[34] describes the results of a two-year randomized controlled trial on biochar adoption and impact among a sample of nearly 1000 farmers in rural Western Kenya. Motivated by provisional findings in the second paper, this project sought not only to test the hypothesis that biochar could be profitable, but to seek out mechanisms by which actors in the sector could stimulate its adoption, starting from the observation that simple profitability is a necessary but insufficient condition for widespread adoption of a novel technology. While the experiment suffers from relatively severe misreporting on the part of respondents, I find substantial impact on crop yields, that biochar's benefit is stronger on farms that yield less, and that biochar has substantial complementarity with inorganic fertilizer. However, uptake was quite low – in spite of substantial agronomic benefit, farmers only purchased biochar when it was very heavily subsidized. While there is some evidence that learning about benefits from having a demonstration plot or learning from the demonstration plots of neighbors increased willingness to pay, demand is insufficient to sustain a profit-making business supplying biochar. Given that biochar is unambiguously profitable, low rates of adoption are socially sub-optimal, particularly when the added value of sequestered carbon is taken into account. From a social cost/benefit perspective, heavy subsidies on biochar are justified, particularly if actors are able to accurately and efficiently target those with the highest potential benefits, which will tend to be those with the most degraded soils.

The dissertation concludes with an essay summarizing lessons learned, and charting the way forward for research with biochar in particular, and technology adoption for sustainable agricultural intensification in general.

Chapter 2

Heterogeneous global crop yield response to biochar: a meta-regression analysis

Abstract
Biochar may contribute to climate change mitigation at negative cost by sequestering photosynthetically-fixed carbon in soil while increasing crop yields. The magnitude of biochar's potential in this regard will depend on crop yield benefits, which have not been well-characterized across different soils and biochars. Using data from 84 studies, we employ meta-analytical, missing data, and semiparametric statistical methods to explain heterogeneity in crop yield responses across different soils, biochars, and agricultural management factors, and then estimate potential changes in yield across different soil environments globally. We find that soil cation exchange capacity and organic carbon were strong predictors of yield response, with low cation exchange and low carbon associated with positive response. We also find that yield response increases over time since initial application, compared to non-biochar controls. High reported soil clay content and low soil pH were weaker predictors of higher yield response. No biochar parameters in our dataset – biochar pH, percentage carbon content, or temperature of pyrolysis – were significant predictors of yield impacts. Projecting our fitted model onto a global soil database, we find the largest potential increases in areas with highly weathered soils, such as those characterizing much of the humid tropics. Richer soils characterizing much of the world's important agricultural areas appear to be less likely to benefit from biochar.[1]

[1]This paper is co-authored with Samuel Abiven (University of Zurich), Simon Jeffery (Wageningen University), and Margaret Torn (Lawrence Berkeley National Laboratory), and has been published in *Environmental Research Letters* 8, no. 4 (2013): 044049.

2.1 Introduction

Biochar – defined as pyrolyzed (charred) biomass applied to soil – has elicited significant interest as a strategy for mitigating climate change, and as an agricultural soil amendment [35–37]. While the long-term persistence of pyrogenic organic material in soil is not well constrained, there is evidence that its residence time can be considerably greater than that of other plant-derived inputs to soil [28, 38, 39]. Production and application of biochar may thereby offer a means of drawing down atmospheric CO_2 concentrations over time by stabilizing dead plant carbon that would otherwise be mineralized more rapidly[40]. Inasmuch as biochar is an effective plant fertilizer, it may offer a "win-win" between climate change mitigation and agricultural production by sequestering carbon, improving soil fertility, and potentially reducing soil non-CO_2 greenhouse gas emissions [41–43].

The upper bound for climate change mitigation through biochar systems has been estimated at approximately 12% of 2010 CO_2-equivalent greenhouse gas emissions per year [36]. Of course, realized mitigation benefits are likely to be lower than this estimated technical potential, with the gap between technological potential and realized benefit depending in large part on the degree to which farmers and other land managers produce and/or apply biochar to the soils that they manage. Given that farmers and land managers are unlikely to utilize biochar in large quantities unless its use is profitable, the degree of biochar technology adoption will in turn be strongly dependent on its costs and benefits – neither of which are well understood.

Increased crop yields are an important part of this benefit. Numerous controlled field and greenhouse experiments have investigated yield response to biochar, and results have been summarized in several reviews [37, 44–47]. Quantitative reviews have estimated average yield benefits from biochar at approximately 10% for crop productivity (encompassing both harvested yields and aboveground biomass production) [44] and approximately 25% for aboveground biomass [45]. Variability of these benefits is high however, ranging between cases where biochar caused a near-failure in plant growth [48] to cases where biochar caused plants to grow where they would otherwise have failed [49, 50]. It seems clear that this variability is mediated by soil properties and properties of the applied biochars. Quantitative reviews have found yield response to be relatively higher in sandy soils, moderately acidic soils, and in response to biochar produced from animal waste[44, 45].

However, quantitative reviews have been hindered by missing and/or inconsistent reporting of soil properties, biochar properties, or other factors which may explain observed plant response. Therefore they have employed univariate meta-analyses of subsets of published data to calculate average responses in under different experimental conditions[44], or fit linear regression models using only those predictor variables that were reported in the majority of available primary studies [45]. While useful as a "first pass" at the data, these approaches may lead to misleading and/or imprecise conclusions stemming respectively from correlation between grouping factors and underlying causes, and low effective sample sizes caused by dropping observations with missing covariate data. The former issue may increase omitted variables bias, while the latter may limit the global significance of these reviews by rendering them unable to use the full set of available studies.

This article addresses these challenges by using statistical methods designed for problems with missing data [51], thereby allowing us to use a more complete set of available studies. First, we seek to infer how biochar's agricultural yield benefit is mediated by varying soil characteristics, biochar characteristics, and management factors. We do so by constructing a semiparametric meta-regression model[52–54] and fitting it to a dataset comprised of 365 observations from 40 studies which compared crop yields in biochar-using treatments to biochar-free controls. Second, we seek to predict where across the globe biochar is likely to have the greatest agricultural benefit. We do so by projecting our fitted model onto a global database of soil properties[55], for a representative biochar at an agriculturally plausible application rate.

2.2 Data and methods

2.2.1 Data

We extracted data on crop and/or biomass yields, biochar properties, and soil properties from 84 studies that examine effects of biochar on plant growth (citations of original studies, as well as full dataset, in SI). Studies were identified using academic search engines, and both peer-reviewed and "grey" literature is included in our dataset. Variables from the original studies were selected for inclusion into our dataset based on consistent availability across original studies, and theoretical importance. Many key drivers of yield could not be included however, particularly soil nutrient contents. We ignore soil nitrogen, phosphorous, potassium, etc., because measures of total content correlate only imperfectly to availability given heterogeneous soil chemistry. Rather than controlling for them explicitly, we account for them indirectly by using random effects (see below).

Our dataset includes studies that reported crop yield, plant biomass production, or both. We use the full set of studies to drive imputation models for the prediction of missing data (described below). We then discard those studies that do not provide measurements of grain, legume, or aboveground non-tree fruit yield. 40 of our 84 studies report crop yields.

Absolute crop yield or plant biomass production is not readily comparable between studies because of heterogeneity in plant species. To make studies comparable, we analyzed data as response ratios [56, 57], defined as the natural logarithm of the ratio of biomass production or crop yield in a given (biochar-incorporating) treatment, to its respective zero-biochar control: $RR \equiv ln\left(\text{Yield}_{\text{treatment}} \div \text{Yield}_{\text{control}}\right)$. This ratio is comparable between diverse studies, while the logarithmic transformation ensures that variability in the ratio's denominator has no greater influence on the metric than variability in its numerator. A response ratio of 0 indicates no change from the control. Response ratio is readily transformed into a percentage relative increase $RI = (e^{RR} - 1) \times 100\%$. If a study applied biochar at multiple rates, we calculate a separate response ratio for each treatment level using a common zero-biochar control mean. We thus construct a dataset of 365 crop yield response ratios, with associated soil and biochar covariates.

Distributions of variables that we analyze, response ratios, and univariate correlations between variables, are presented in supplementary figures S1, S2, and S3. Our data contains a broad range of soil properties, biochar properties, and response ratios, though with some gaps and anomalies. For example, no studies in our dataset used biochar produced between 650 and 900°C. Furthermore, there was no strong univariate relationship between percentage carbon and pyrolysis temperature. This was somewhat unexpected, given laboratory studies showing increasing carbon content as a function of pyrolysis temperature[58], though it is possible that this simple correlation is driven by other correlated factors, such as feedstock or pyrolysis technology. Increasing pyrolysis temperature was correlated with increasing pH [59], while increasing pH was associated with lower carbon content. Based on bivariate comparisons of available data, neither the response ratio nor its variance were well-explained by any of the variables singly (figures S1, S2, and S3), though there are suggestive correlations between response ratio and soil pH and CEC. This lack of strong univariate explanatory power was a key motivation for our use of the multivariate methods. Finally, we classify the biochars used into one of three categories; "woody," "nonwood," and "manure." "Woody" biochars are those made from wood, "nonwood" biochars are those made from plant matter other than wood (mainly crop residues), and "manure" biochars include any that are derived from animal waste. 18 studies in our crop yield dataset use nonwood chars, 18 use wood chars, and 8 use manure chars. For variables that we include in our analysis that were missing in primary studies, corresponding authors were contacted with requests for those data. Nevertheless, we lack observations on one or more variables from the majority of the studies in our dataset (figure S4).

2.2.2 Methods

Our dataset and context have a number of distinctive features that have shaped the analytical methods that we use. First, our dataset is comprised of studies, rather than primary data. We therefore draw upon "meta-regression" techniques[60, 61], in which observed effects are modeled as a function of a study-level random effect and a set of explanatory fixed effects. We specify random effects at the level of unique soil-biochar combination. This is a slightly lower level than study-level, and is motivated by the fact that several studies investigate several different soil-biochar combinations. It is nonetheless a higher level than a random effect at the common-control experiment level; given that there are 230 of these, with many singletons, specification of random effects at this level would adversely affect confidence intervals [62] while also over-fitting the model. Observations are weighted such that each common-control experiment in our dataset is given equal weight, i.e. such that common-control experiments applying biochar at several different application rates do not influence our estimates any more than studies examining fewer application rates. As is standard practice in meta-regression, we further weight our observations by the inverse of the variance of the response ratio – a practice aimed at addressing publication bias and up-weighting studies whose effects are more precisely estimated[60, 61]. Because our distribution of response ratio variances is highly skewed and in all cases between zero and one, we weight by the logarithm of the inverse of the variance.

Second, many of our observations contain missing data along one or more covariates of interest (figure S4). We use multiple imputation[51, 63] to draw inference using observed data from partially-missing vectors of observations. Specifically, we construct multiple imputations via chained equations (MICE) [64, 65], with a relatively large number of imputations ($M = 200$) chosen for numerical stability and result replicability. Details of our imputation modeling and R code for implementation are available in SI. Briefly, the MICE algorithm operates as follows: It begins by filling in missing values with arbitrary values. It then cycles through each variable and constructs a regression model to fit the partially-observed dependent variable to a partially-observed and partially imputed dataset comprised of the other variables. These models are then used to update the initial arbitrary imputations, in the form of a draw from the fitted regression model's predictive distribution. We cycle through the set of variables many times until our imputations approximately converge in distribution, saving the values at the final step. We repeat this entire process once for each of the $M = 200$ imputations, generating 200 slightly-different datasets with missing values "filled in" with plausible guesses as to what the missing values might have been. The 200 fitted models are then combined according to Rubin's rules [51]; averaging coefficient vectors, averaging variance-covariance matrices, and adding a non-negative correction to variance-covariance matrices that is inversely proportional to the predictive ability of our imputation models. The effect is to widen confidence intervals where there is either a lot of missing data or where missing values of data are poorly predicted by observed data.

Third, we have little prior justification for assuming linear relationships between our observed variables and crop yield response ratio, nor do we have strong *a priori* knowledge of appropriate functional forms to use. We therefore use semi-parametric techniques to estimate smooth functions mapping our independent variables to RR, employing generalized additive models[66] fit with the `mgcv` package in R [52, 53].

Finally, because our dependent variable is the logarithm of the ratio of yields in a biochar-incorporating treatment over a zero-biochar control, its expectation should go to zero as the biochar application rate goes to zero. We are therefore unable to simply include the biochar application rate as an additive term in a statistical model, as such a specification would generate non-zero estimates of $\hat{RR}$ at nil biochar application. We therefore employ a two-step procedure: we first estimate slopes mapping RR to biochar application rate, allowing for heterogeneity in slope coefficients by common-control experiment. Several specifications for estimating these slopes were estimated – the best-fitting model fit a fixed slope to the logarithm of the biochar application rate (figure S8). Other specifications tested includes linear slopes and random coefficients models (figures S6-S12). See SI for further information. From this fitted model, we calculate $\tilde{RR}_{is,BC=3\text{Mg ha}^{-1}}$, which is an estimate of the response ratio for a given study at 3 Mg biochar ha^{-1}. We chose 3 Mg ha^{-1} because relatively low biochar application rates are likely to be more agronomically realistic than the often-higher rates used in research studies.

To explain heterogeneity in response to biochar, we then fit a statistical model to each of our 200 sets of estimated $\tilde{RR}_{is,BC=3\text{Mg ha}^{-1}}$ and associated partially-imputed soil and biochar covariates. Our model is specified in equation 2.1:

$$\begin{aligned}
&\tilde{RR}_{is,BC=3\text{Mg ha}^{-1}} = \\
&\quad \alpha + \tau_s + \beta_1(\text{Pot trial==1}) + \beta_{2-4}(\text{Seasons}) \\
&\quad +\beta_{5-7}(\text{feedstock types}) + \beta_{6-17}(\text{Crop}) \\
&\quad +f(\text{N app. rate}) + f(\text{Py. temp.}) + f(\text{BC \% carbon}) \\
&\quad +f(\text{BC pH}) + f(\text{Soil CEC}) + f(\text{Soil pH}) \\
&\quad +f(\text{Soil org. C}) + f(\text{Soil \% clay}) + \epsilon_{is}
\end{aligned} \tag{2.1}$$

β terms indicate parametric coefficients. $f()$ terms indicate non-parametric smooth functions, represented by cubic regression splines with 10 knots spaced evenly over deciles of the unique values of each variable. In addition, in supporting information we present a more complex model involving a number of interaction terms, designed to better represent interdependence among the drivers of plant growth (SI section 7). While doing so improves model fit and AIC slightly, results are qualitatively similar and individual model terms lose statistical significance due to collinearity.

The fitted model was then projected onto a global dataset of soil properties[55], after masking out non-agricultural areas using data on global cropland extent[67]. This dataset provides 0.5°-resolution data on numerous soil properties, including all of the soil variables in our model. While soil properties often vary dramatically in a $0.5° \times 0.5°$ gridcell, the dataset facilitates our exploration of the average effect of biochar across large geographic areas, recognizing that this spatial averaging may mask local heterogeneity. In addition, the dataset provides information on heterogeneity within each gridcell; each gridcell is comprised of a number of different soil series and a fractional spatial extent within the gridcell. For the purpose of creating a representative soil for each gridcell, weighted averages for each gridcell were taken, with weights corresponding to the spatial extent of each within-gridcell soil series. Estimates of relative increase for each gridcell were mapped by calculating the expected response ratio for each gridcell as $\hat{\tilde{RR}}_{xy,BC=3\text{Mg ha}^{-1}} = \hat{f}(X_{xy})$, where $\hat{f}()$ indicates our fitted model 2.1, and X indicates spatially averaged soil properties for gridcells indexed by latitude x and longitude y.

A more detailed description of our methodology, along with R code for implementation of our entire analysis is available in SI.

2.3 Results

We estimate an average crop yield increase of approximately 10% for 3 Mg ha^{-1} biochar addition in the first year after application (figure 1). Variability in this response is high, ranging from cases where biochar reduced yields to cases with large relative increases (commonly from cases with near-failure in zero-biochar controls).

Soil properties were the best predictors of this variability (figure 2), with low cation exchange capacity and low organic carbon content associated with positive yield response. Low soil pH and high soil clay content[2] were weakly (and non-significantly) associated with positive yield response. Yield response to biochar increased significantly over time, by approximately 0.068 response ratio units in the second season after application, to approximately 0.117 response ratio units in the fourth season after application (corresponding to 7.0% and 12.3% percentage point relative increases in crop yields, respectively). We found little evidence that plant response to biochar is mediated by nitrogen additions to soil[3].

Yield response was invariant to biochar type as parameterized by biochar pH, carbon content, and pyrolysis temperature. The benefit of manure biochar was higher on average than that of wood or non-wood biochar, but the effect was highly variable and not statistically significant. When fitting model 2.1 without observations from manure chars, or with only wood chars or only nonwood chars (not shown), results were nearly identical to results from models with all chars included together, though statistical confidence decreases due to smaller sample sizes. Similarly, while there was heterogeneity in response among crop types[4] and experiment type (field vs greenhouse conditions), these differences were noisy and not significant.

Since soil properties predict yield response, spatially explicit analysis of where biochar is likely to have a greater or lesser benefit is possible using global databases of soil properties.

[2]We note that our definition of "clay content" is simply the clay content presented by authors of primary studies. Many primary studies do not differentiate between clay minerals and clay-sized particles, such as aluminum and iron oxides. Figure S18 presents an estimate of the effect of clay content (as defined by authors of primary studies) as mediated by soil CEC. We find suggestive evidence that high clay content is a driver of positive response to biochar particularly when cation exchange capacity is low.

[3]This result reflects the respecification of N application rate as a linear term, because its estimation as a smooth function led to highly implausible estimates of extreme nonlinearity. Nonparametric estimates suggested sharply negative response ratios at low N applications and sharply positive ones near 100 kg N ha^{-1}, before leveling to a modestly increasing function beyond 150 kg N ha^{-1}. We therefore re-parameterized it as a linear parametric term. As such, it loses statistical significance.

[4]Robustness checks are provided in SI (figure S15) exploring heterogeneity in response to our predictor variables by crop type. Only very modest heterogeneity was found, supporting our specification of differences in response by crop as fixed indicator variables uninteracted with soil or biochar parameters.

Figure 2 projects our fitted model onto the ISRIC-WISE database of global derived soil properties [55] for a non-wood biochar having median values of all biochar variables, and applying a dataset-median amount of nitrogen. Supplementary figure S14 gives the same map, with associated standard errors of fitted values. Values presented in figure 3 have the interpretation as our model's estimates of average plant response to biochar solely as a function of soil and biochar properties. Other factors that may influence yield response to biochar – such as heterogenetiy between crop species, or differing biochar efficacy across different climates – may introduce bias in our effect estimates in model 2.1 or in our predictions in figure 3 if they are correlated with variables present in our model. In addition, our map abstracts from within-gridcell heterogeneity, and only presents spatially-averaged estimates.

Given the global distribution of soil properties that we include in our model, our fitted model implies positive yield response over much of Sub-Saharan Africa, parts of South America, Southeast Asia, and southeastern North America. These areas are largely coincident with areas of highly weathered soils, generally in the tropics or subtropics. Our fitted model also implies that response in the the belt to the north of Eastern Europe's chernozems would likely be positive. Yield response is predicted to be most negative in organic soils such as those in Indonesia, northern Eurasia and North America. Notably, our model implies that yield response may be weak and/or negative in many of the world's most important grain-producing areas, such as the Eurasian chernozems or central North American mollisols. Implications for South Asian vertisols, and in much of the North American corn belt near the great lakes, are predicted to be small to negative.

While our dataset gives good coverage over our parameter space (Figures S1-S3), geographic coverage is relatively poor in many regions. No studies have reported yield response to biochar in central North America, Eastern Europe/Eurasia, the Atlantic coast of South America, the African Sahel, or the South Asian peninsula. While soil properties characterizing these regions are represented in our dataset, our predictive maps should be viewed with caution where geographically out-of-sample predictions are being made.

2.4 Discussion

Our findings differ from previous meta-analyses of plant response to biochar, which used qualitative methods, univariate comparisons or complete-case regression analysis [37, 44–47]. Those studies found strong correlations between response ratio and soil pH, benefits in sandier soils, and stronger benefits in response to animal-derived biochars than plant-derived biochars. We find only suggestive evidence that soil pH influences response ratio, and suggestive evidence that soils with low reported clay content can expect weaker benefits to biochar application. While we estimate that response is higher in animal derived biochars, the result is not significant. In contrast, our results emphasize the roles of soil cation exchange capacity and soil organic carbon content, with weaker emphasis on soil pH and opposite results for texture (as defined by reported soil clay content).

Our estimates and ultimately our projection maps are consistent with biochar's benefits being maximized on weathered and degraded soils, which tend to have low cation exchange capacity, low soil organic carbon, low pH, and relatively non-reactive clay mineralogy. Somewhat surprisingly, we find little association between response ratio and biochar properties such as percentage carbon, pyrolysis temperature, or pH. Indeed, this lack of association (which is also apparent in bivariate scatterplots in figure S2) suggests that if properties of biochars themselves are relevant in determining yield outcomes, the variables which we inclue in our model are of either unimportant, or potentially important only in mediating the influence of other variables. Explaining mechanisms by which different biochars influence yield outcomes remains an area for future research.

One of the more interesting and novel findings of our analysis is that biochar's yield benefits significantly increase over time, which was found individually in several studies in our dataset [49, 50, 68–71]. Many of these found slower rates of yield decline under continuous cultivation using biochar[49, 71], where both control and treatment yields declined after the initial season of implementation. This distinction between absolute increases versus slower decreases is important, and deserves further attention in future research. Response heterogeneity over time also warrants further research – our rather simple parametrization of time periods as un-interacted dummy variables captures only mean response over time[5]. However, plant response to biochar over time has important implications the economics of its use, and thereby its ability to contribute to climate change mitigation and improved agricultural productivity. Many biochar systems are costly in terms of labor or investment. Mechanisms underlying sustained/increased yield over time – along with improved quantitative understanding of biochar's retention [73–75] and decomposition within [28] soil – should improve understanding of the conditions under which biochar might be economically viable.

Finally, we present a few caveats to the interpretation of our results. First, we emphasize that our model presents average effects of our specified dependent variables on response ratio. While we find only small evidence for heterogeneous effects by crop or by biochar type, we are unable to exclude their presence. Second, our statistical approach will inevitably create misspecification bias: the predictor variables in our dataset are highly interactive in soil, yet we use them to fit an additively separable statistical model.[6] However, even when investigating a somewhat richer model with several interaction terms (equation S10, figures S16-S20), results are largely consistent with this more simple model. We therefore favor the more parsimonious model for ease of interpretation. Furthermore, given that results from model S10 are generally somewhat larger, we

[5] In our dataset, the largest increases over control (>150%) beyond year 1 are from poor soils in Brazil [49] and Zambia [50]. Multi-year responses in other soils were generally in the 10-40% range. In one study that grew a mix of vegetables on a high-CEC, alkaline soil [72], plant response to biochar actually declined over time.

[6] To name a few interactions, cation exchange capacity is determined both by clay mineralogy, soil organic matter, and pH. Biochar's surface charge is pH-dependent [76], and it is known to become encapsulated by clay particles in aggregates [77]. While its own pH tends to be high, the degree to which soil pH is affected by biochar will depend on buffering capacity.

favor the more parsimonious version as a more conservative lower bound on potential outcomes. Finally, predictions made from an imperfect model will inevitably propagate error. However, as the number of biochar experiments increases with time, statistical models specified based on a more detailed representation of biochar's dynamics in soil would represent a useful avenue for future research.

2.5 Conclusion

Our estimates suggest that biochar has a substantial and specific agroecological niche – on poor soils characterized by low cation exchange capacity, little soil organic carbon, and perhaps with lower pH and heavier textures. These characteristics describe many of the world's lowest-potential agricultural areas, which are predominantly found in the humid tropics. Given that many of the world's poorest agricultural soils exist in locations that are coincident with high rural poverty [4], our analysis invites further research on biochar's potential to play a role in agricultural development. Given that carbon mineralization from dead plant matter is generally faster in warm and moist conditions, biochar may represent a more effective means of carbon sequestration in the soils where it has the most agricultural benefit, than in soils where it has less benefit. This alignment however would depend on the degree to which biochar mineralization is dependent on environmental characteristics versus on its intrinsic chemistry, which for biochar remains poorly understood [28, 78].

Although areas for future research remain, given economically viable production technologies, technology adoption by potential beneficiaries, and an enabling policy environment, biochar systems appear to have potential to make a serious contribution to world food production, particularly in agricultural areas that are currently somewhat marginal in terms of global food production. Given increasing global food demand, this represents an important opportunity. While the economic viability of large-scale biochar systems remains to be demonstrated [79], substantial agronomic benefit over multiple years at relatively low application rates in certain regions indicates the possibility that some level of climate change mitigation via biochar systems could be achieved at negative cost.

2.6 Figures

FIGURE 2.1: Figure 1: Kernel density plot of estimates of relative increase in crop yields against zero biochar controls in the first harvest after application, in response to 3 Mg ha^{-1} biochar. Dotted red line indicates the sample mean.

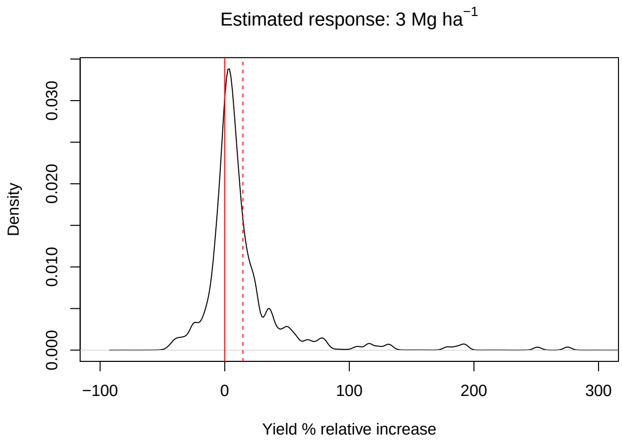


FIGURE 2.2: Figure 2: Estimated nonparametric smooth functions and parametric slope coefficients for factor variables, from model 2.1. Printed p-values are based on F-tests that $\hat{f}(.) = 0$ [80]. Shaded bands or dashed blue lines indicate 95% confidence intervals for the mean effect of each term. R^2 averages .81 across 200 imputations. Figure S13 gives distributions of residuals (from each of the 200 fitted imputation models) for each datapoint, by study. The model is significantly different from a null model (with only random effects and no covariates) at $p = .003$.

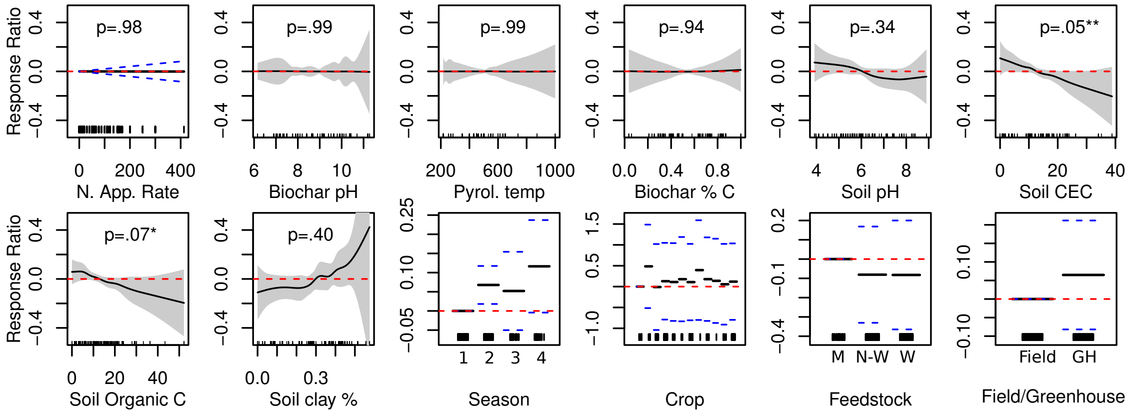
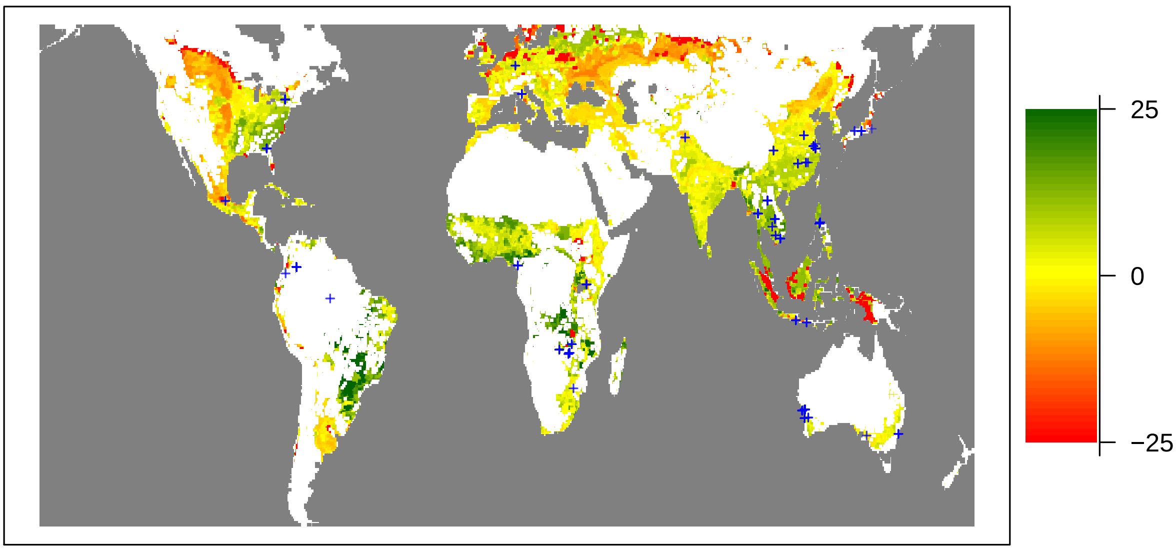


FIGURE 2.3: Figure 3: Model predictions for a median non-wood biochar applied to maize, projected onto spatially-averaged soil property dataset from Batjes (2005)[55], in units of percentage relative increase. Areas with < 5% of land in agriculture are masked using agricultural area extent data from Ramankutty [67]. Blue crosses give locations of studies used to fit the model.



Chapter 3

Biochar profitability and carbon prices: Evidence from Kenya and Vietnam

Abstract

Biochar can increase agricultural yields while sequestering carbon in soil, thereby contributing to climate change mitigation. However, it is unclear whether biochar requires a carbon subsidy to render it economically viable. Using recall-based datasets from Kenyan and Vietnamese smallholder farmers, we estimate yields as a function of biochar and fertilizer use. We find an positive association between biochar use and average yields in Kenya, but no correlation in Vietnam. We then use these estimates to calculate optimal input mixes under hypothetical biochar and carbon prices, given heterogeneity in response both to biochar and fertilizer, and heterogeneous budget constraints. In Kenya, we find that biochar is more-likely-than-not to be profitable to adopt for 23% of our sample if unsubsidized and available at its current sale price of around $188/ton, while a hypothetical carbon subsidy of $100/ton CO_2e increases this proportion to 47%, though these proportions are not different from zero at 95% confidence. Because of limited short-term complementarity between biochar and inorganic fertilizer, we estimate that biochar adoption would change profits little, given budget constraints for agricultural inputs. We conclude that carbon subsidies may have a marginal impact on biochar's profitability in Western Kenya, but that further research is needed to improve the precision of these estimates, extend them to account for any longer-term changes in soil characteristics that might impact biochar's profitability, and account for any potential biases stemming from time-varying variables that not measured or modeled in the context of this study.[1]

[1]This paper is co-authored with Abigail Clare (University of Edinburgh) and Alessandro de Pinto (International Food Policy Research Institute), and has been submitted for publication in a peer-reviewed journal.

3.1 Introduction

Efforts to increase organic carbon content in agricultural soils are growing, particularly in the context of programs aiming to mitigate climate change through soil carbon sequestration[79, 81]. One method that is gaining increasing attention is the stabilization of crop residues through pyrolysis – producing charred material termed "biochar" when used as a soil amendment [35, 82]. The carbon in biochar (typically 50-90% by mass) has a substantially longer residence time than carbon from similar un-charred plant material, and as such may represent a means of sequestering carbon in soils. While biochar's residence time in soil remains uncertain, most estimates suggest centennial or millennial timescales [28, 38, 83], as opposed to sub-annual to decadal timescales for un-charred plant inputs [84]. Similar to other forms of soil organic matter (SOM), biochar can improve soil fertility across a number of dimensions – improving water holding capacity and soil structure, while reducing acidity and improving fertilizer-use efficiency [37, 47]. In areas with weathered soils or low/depleted levels of SOM, biochar may not only improve soil fertility in ways similar to the addition of SOM, but potentially do so in a way that is durable over time [32, 49]. Most famously, *terra preta de indio* soils in the Amazon basin, to which pre-Columbian populations added considerable quantities of biochar, have retained high fertility since being abandoned several centuries ago, and much of the carbon in these soils has been dated at >1000 years old [85]. These intriguing agronomic and biogeochemical properties are leading to increasingly applied research, as well as a fledgling industry[86].

Recent interest in biochar has largely been catalyzed by the prospect that it might mitigate climate change while increasing yields or reducing production costs. Meta-analyses of crop yield response to biochar applications have found average crop yield increases of approximately 10% over controls, and substantially higher benefits in highly-weathered and/or degraded soils [32]. Biochar may also affect other soil properties, leading to reduced nutrient leaching (thus improving fertilizer use efficiency)[42, 87–90], and lowered CH_4 and N_2O emissions from soil [42, 43, 91].

However, the profitability of biochar usage remains understudied and it is unclear what role it might play in climate change mitigation. Extant studies on the subject have employed basic economic modeling to assess *ex ante* biochar's profitability [92–96], or modeled optimal pyrolysis temperatures given exogenous prices for biochar [97], or inferred break-even prices for biochar from single agronomic trials [98]. In particular, Pratt & Moran[99] developed marginal abatement cost curves for biochar, but their results were highly sensitive to both carbon price assumptions and assumptions about agricultural benefit. Most of these studies have focused primarily on biochar utilization in developed countries (where it may have lower potential agronomic benefits on average given prevalent soil types in the global north [32]. An exception is Galgani et al.[100] who use a life-cycle-assessment framework to model return-on-investment as a function of carbon subsidies, in the context of biochar production from municipal solid waste in

Ghana. They find that the profitability of biochar production is highly sensitive to monetary compensation for carbon sequestration, becoming viable at 30 euros per ton CO_2e, though their result is driven by the assumption of an baseline scenario in which that solid waste would generate substantial greenhouse gas emissions as a result of landfilling. No extant studies have included any empirical analysis of benefits for biochar users largely because little data on biochar-using farmers existed.

Our approach is distinct from the controlled agronomic experiments and theoretical modeling studies that have comprised the literature to date. This paper presents the first empirical analysis of returns to biochar use among smallholder farmers, using novel recall-based datasets from early adopters in Kenya and Vietnam.

Our two regions provide a substantially different economic and environmental context than previous studies, and in turn are quite different from one another. Our western Kenya dataset is based on a survey of subsistence or semi-subsistence rainfed maize farmers. By contrast, our Vietnam dataset is based on a survey of farmers practicing irrigated, intensive rice farming with relatively high levels of inputs. Both of these cases represent examples of global challenges in agricultural development. The Kenya case can be compared to much of the rest of sub-Saharan Africa, where fertilizer use remains low and soil nutrient mining is widespread[9, 10, 101]. This can result in low productivity, land degradation, and potentially expansion of the agricultural frontier into environmentally sensitive areas. While many advocate an increase in the use of inorganic fertilizers, the profitability of these inputs is often conditional on soil fertility itself, and levels of soil organic matter in particular [6, 7]. As such, technologies that use locally-available resources to enhance natural soil fertility and increase the efficiency of inputs such as inorganic fertilizer may contribute to agricultural development.

The Vietnam case represents a different problem; agriculture in much of east and southeast Asia is extremely intensive and application rates of inorganic fertilizers are very high. Not only does high use of inorganic fertilizers have negative environmental consequences [102–104], there is evidence that it is often used to the point where marginal private costs exceed marginal private benefits [105–107]. In this context, it is hypothesized that the use of biochar may improve fertilizer efficiency, thereby improving economic returns and reducing environmental impacts of agriculture in the region.

Though our results are weakened by potential measurement error stemming from recall-based data, together with the possibility of omitted variables bias, we find both crop yield increases (in Kenya), and that biochar led to reductions in inorganic fertilizer expenditure. In Kenya, we find that biochar adoption led to 17% average increases in yield and an average decrease in fertilizer expenditure of about 26%, along with considerable heterogeneity in response to fertilizer, echoing similar results in the region[25]. However, we find little evidence of complementarity between biochar and fertilizer on average. This is somewhat in contrast to results from the region showing that soil carbon improves response to fertilizer, implying either that biochar does not act in a manner similar to other forms of soil carbon, or that any effect must be longer-term – not manifesting after 1-3 years. Our estimates imply that biochar would be profitable for about 23% of our sample unsubsidized, increasing to 47% assuming a carbon price of \$100/ton CO_2e. Given uncertainty in the production function however, neither these proportions nor the change in these proportions is distinguishable from zero at 95% confidence. Our estimates also imply however that this adoption would change optimized profits and fertilizer expenditure little, given little complementarity between biochar and fertilizer in the short run. In Vietnam, we find little evidence that biochar adoption improved yields or response to fertilizer, though we find evidence in some specifications that farmers who adopted biochar reduced their use of fertilizer. As such, we have weak evidence that biochar enabled farmers to maintain yields while reducing fertilizer expenditure, thereby improving the efficiency of fertilizer use.

Altogether, we find that biochar may have substantial agronomic utility in some settings, but we find limited evidence that it may stimulate intensification and limited evidence that it might play a substantial role in climate change mitigation. While further research is needed to improve the precision of our estimates, we argue that these findings justify that research.

3.2 Approach, data, and methods

We seek to address two distinct, yet related, questions. First, we seek to identify the causal effects of biochar adoption on yields. In Kenya, where our data is only comprised of biochar adopters, this is an average treatment effect on the treated (ATT). In Vietnam, where our data includes non-adopting farmers, this is an average treatment effect (ATE). Given that farmers will not generally hold all else constant when adopting a new technology, these effects represent both direct agronomic effects of biochar on crop yields, as well as effects of changes to other inputs that farmers may have made when they began to use biochar.

Second, we seek to characterize the effect of carbon prices on biochar's profitability. We do so by first estimating a production function relating inputs to yields, and then calculate optimal mixtures of biochar and fertilizer across a gradient of carbon prices, from \$0-\$100/ton CO_2e[2]. This allows us to estimate a change in carbon prices would affect the proportion of farmers for whom it would be profitable to adopt biochar.

[2]The maximum price we use, \$100/ton CO_2e, is substantially higher than can be found in any extant trading schemes. However, it is lower than the \$43/Mg C (\$158/Mg CO_2e) estimate of the social cost of carbon taken from a survey of published estimates[108].

3.2.1 Data

3.2.1.1 Kenya

Our Kenya dataset was collected from early biochar adopters in Bungoma county and parts of northern Kakamega county, western Kenya, in June and July of 2011. Agriculture in the region is dominated by maize and sugar cane production. The former is grown mostly for consumption, while the latter is a cash crop. The region receives about 1600mm rainfall annually, distributed bimodally in a long and a short rainy season, with the former lasting approximately March to June and the latter from August to November. Most farmers plant two maize crops per year, and yields in the short rains are typically half of those from the long rains. Soils in the region are relatively weathered. Population density is higher than that in much of the rest of rural Africa, averaging 100-250 persons/km. Farm sizes average below 1 hectare, and fallowing – once common in the area – is now rarely practiced due to land scarcity.

In these two districts, approximately 700-800 farmers had been trained on the use of biochar by the African Christians Organization Network (ACON), a small Kenyan NGO. ACON began these trainings in early 2008, after good results were observed on small plots. ACON typically targeted self-organized community groups, who established biochar demonstration plots under ACON's supervision. After observing good results on these demonstration plots, training participants often moved on to use biochar on their own land. New trainings in new areas typically occurred after friends or relations of previously-trained farmers learned about ACON's work and requested trainings for themselves. Most farmers in our sample used biochar from one of three sources: (1) collection of charcoal fines (small pieces and dust) from charcoal sellers in the area, (2) partially-combusted residues of home cooking fires, and (3) home production of cooking charcoal (commonly produced for sale), and utilization of the resultant dust and fine particles as biochar[3]. None of these sources is readily scalable. The first and second sources are limited inasmuch as they represent waste products of other activities, while the third source is constrained both by Kenyan law, and by the fact that tree-cover is increasingly limited to trees planted and actively managed for timber and other uses. Farmers using biochar report that they are constrained by supply – many would like to use more.

More scalable solutions for producing biochar – whether centralized facilities or farm-scale technologies – were not present in the region at the time that we collected these data. This has changed in the intervening period – in 2013, a project operating in the area began producing and selling biochar to farmers, at a price of KSh400/25kg sack (about $4.5) – an amount sufficient for about 1/8th acre (∼1/20th of a hectare) – from sugar

[3]There is no formal difference between cooking charcoal (generally produced from wood) and biochar. However, some forms of charcoal are not suitable for fuel, such as charred maize stover or charred sugarcane leaves. No farmers reported purchasing wood charcoal to pulverize and add to their soil, suggesting that farmers perceive wood charcoal to have a higher use value as fuel than as a soil amendment. While it is possible to make fuel briquettes from chars made from cellulosic feedstocks such as sugarcane leaves or maize stalks, this practice was not apparent in the area at the time that we conducted our research.

cane residues, using simple kilns constructed from modified oil drums. We use these parameters when modeling optimal biochar adoption decisions.

We surveyed 361 biochar-using farm households in June and July of 2011. While ACON had conducted trainings throughout Kenya and into Uganda, we restricted our survey to a census of participating households within the borders of Bungoma and Matungu districts, which adjoin and surround Bungoma town. After removing observations with missing data on adoption timing or other serious abnormalities, we are left with a sample of 302 farm households.Surveys focused on maize yields and fertilizer expenditures – all on a per-acre basis – along with basic demographic information.

We rely on self-reported data from the long rain harvest of 2008 to the season immediately preceding our survey. Use of recall data may affect our results in one of two ways: either through errors in memory or systematic untruthfulness. We have no reason to suspect the latter, though the former could either be a source of classical measurement error or a source of bias if farmers may systematically recall good or bad experiences differently.[4]

Because of the likelihood that the farmers who chose to use biochar are systematically different from those who did not, we did not survey farmers who had not themselves chosen to use biochar. Instead of constructing a counterfactual from farmers who never chose to adopt biochar, we exploit heterogeneity in adoption timing between biochar adopters.[5]

These surveys were conducted on an extreemly limited budget, and as such, they have several shortcomings. First, we lack data on several potentially-important omitted variables. While the longitudinal structure of the data that we collect allows us to control for all stable covariates, the possibility that there may be important uncaptured varaibles

[4]In a test of the accuracy of recall-based data from coastal India, de Nicola and Giné[109] find evidence that particularly favorable past outcomes are more likely to be reported accurately, while less favorable outcomes tend to be forgotten. If this same dynamic holds in our data – probability of upward bias increasing with distance from the present – then our estimates would tend to be biased downwards, because upward bias in past outcomes would reduce the difference between pre- and post-adoption outcomes. Alternatively, unfavorable past outcomes might be more cognitively salient, influencing perception of normal previous seasons and biasing our estimates upwards. While we cannot be certain of either the existence or the sign of any potential recall bias, we take some reassurance in the fact that our data on maize yields roughly matches aggregate trends for Kenya – while our sample produced an average of 2.4, 2.7, and 3.3 tons/ha in the combined long and short rainy seasons of 2008-2010, the country as a whole produced 2.3, 2.4, and 3.4 million tons.

[5]As such, our counterfactual group is comprised of biochar adopters *during the periods before they adopted biochar*. Our estimates will be interpretable as average treatment effects on the treated (ATT) – the effects on farmers who themselves choose to adopt – rather than a treatment effect for an average farmer (ATE). Given that our data is longitudinal, our estimate of the effect of biochar adoption is derived from time periods in which the transition from non-adoption to adoption is incomplete, and causal interpretation of estimates requires that differences between adoption groups be purely cross-sectional, rather than time-varying (i.e.: an assumption of "common trends"). Data from later periods – by which time almost all in our sample had adopted biochar – plays no role in identifying the treatment effect, and merely serves to improve statistical precision.

that vary with time cannot be excluded. These might include seed varieties, labor allocation/expenditure, intercropping, and other inputs like tractors or fertilizer. We do not know the extent to which their omission will bias our estimates, though we consider their potential impact in the discussion. Second, because farmers themselves were typically unable to approximate how much biochar they used, our enumerators resorted to a form of estimation which we felt to be highly-leading of the respondents. We therefore discontinued questions on biochar quantity, and in all models represent biochar usage with a dummy variable. Thus, our estimates will carry the interpretation of the average change observed after using biochar in whatever unknown quantity farmers chose to use it.

Summary statistics by adoption season are given in table 3.2. The large majority of farmers began using biochar between 2009 and 2010. It is apparent that there is some heterogeneity in groups by adoption season. Long rains 2010 adopters had slightly more non-farm income and were slightly more likely to be male, and those few who adopted biochar ahead of the first and last seasons in our dataset tended to have lower yields and more land allocated to maize.[6]

Given that our data is longitudinal, we will control for farmer-specific fixed effects in all models. As such, causal interpretation of our estimates will require that unobserved time-varying factors affecting yield are balanced across adoption groups. Given that this is fundamentally untestable, we are forced to take it on as an assumption. However, we provide a partial check against its violation by directly checking for differential trends in our outcome variables by adoption group. We do so by first deseasonalizing yields by regressing them on season dummies, and then calculating a trend for farmer i in time t as the difference between one season's de-seasoalized yield and its lag. We then fit models of the form $trend_{i,t\neq T} = \mathbf{A}'\beta + \epsilon$ to subsets of the data for each season (excluding current-season adopters), where $\mathbf{A}$ are factor variables for the season of initial adoption. Results are given in table 3.3 – we find no evidence of differential trends either pre-adoption or post-adoption for any of our adoption cohorts. This is consistent with our understanding of the mechanism by which ACON diffused biochar in the area – conducting trainings in new areas after good results spread by word-of-mouth, rather than adopting spreading from early adopters to later adopters within single areas, which might be indicative of non-ignorable selection into treatment.

Causal interpretation of our estimates is also threatened if different adoption cohorts face systematically different returns to adoption. If this were the case, then farmers with higher potential returns to adoption would be likely to select into adoption sooner than those with lower potential returns, and the difference that we estimate would be substantially higher than the difference that we would observe if the cohorts switched

[6] Our panel is not entirely balanced, because farmers sometimes didn't farm for particular seasons. This can stem from a number of factors, including illness, lack of money to purchase inputs, or a desire to farm only during the long rain season rather than the short. If the decision not to farm is related to time-varying unobserved shocks and also related to adoption timing and to outcome variables, then treatment effect estimates will be biased. However, results from Poisson and OLS regressions of the number of seasons not farmed on adoption season suggest little to no correlation between these variables.

the order in which they received biochar, and also higher than they would be if selection of adoption timing was effectively random. Therefore, we directly compare potential outcomes by simply comparing the change in (deseasonalized) yields in each cohort's first season of adoption, via a regression of the form $trend_{i,t=T} = \mathbf{A}'\beta + \epsilon$. Here, we find that season 4 adopters experienced a statistically significantly larger jump in yields on biochar adoption than other cohorts. We therefore exclude season 4 adopters from our sample hereafter.

3.2.1.2 Vietnam

Our dataset of early biochar adopters in Vietnam is comprised of farmers trained between 2009 and 2012 in Ha Noi, Hung Yen, Hai Duong, and Thai Nguyen provinces of northern Vietnam, and is based on a household survey conducted in July and August of 2013. Paddy rice is the region's dominant crop, though horticulture and animal production are also important. The region receives plentiful annual rainfall (approx 1700mm/year), and most farmers plant and harvest rice twice per year, in the spring/summer and summer/fall growing seasons.

Biochar trainings were initiated by the Institute for Agricultural Environment (IAE), a research institute under Vietnam's Ministry of Agriculture and Rural Development. Much of the impetus for the initiation of these trainings stemmed from seasonal burning of rice straw and rice husk by farmers near urban Hanoi; given more crop residue than was needed for animal bedding, animal fodder, and other similar uses, farmers burnt large quantities of it, leading to air quality problems in surrounding areas. In this context, biochar was seen as a means of valorizing a waste product while mitigating a local environmental problem. Trainings began in late 2010 in Ha Noi province, in late 2011 in Hai Duong province, and in mid 2012 in Hung Yen province. We also have data from Thai Nguyen province, where biochar dissemination was conducted by CARE Vietnam and the University of Thai Nguyen since early 2009, in collaboration with IAE.

As with our Kenyan data, farmers self-selected into biochar outreach activities. Many were motivated by a desire to enhance soil fertility, while others were motivated by a desire to find alternatives to the (illegal) burning of rice residue. Many farmers produce biochar via simple equipment developed and provided freely by IAE, while others produced biochar by piling feedstock, setting it on fire, and covering it with soil in order to choke oxygen supply.

Our Vietnam sample includes 1545 observations of 296 farmers in Ha Noi and Hai Duong provinces, who were selected for the survey at random from lists of households in the areas where IAE had conducted biochar trainings. We do not include data on farmers from Hung Yen or Thai Nguyen in our analysis, because all (none) of the farmers who used biochar in Thai Nguyen (Hung Yen) had begun using by the first remembered period in our dataset, leaving us unable to use pre-adoption farmers as counterfactuals for themselves post-adoption. Because farmers themselves were typically unable to approximate how much biochar they used, we specify biochar usage as a dummy variable. In this case we have data both on biochar adopters (N = 221) and non-adopters (N = 75), and we use both adoption and adoption timing to create counterfactuals against which we can measure biochar's impact. As with the Kenyan data, the Vietnamese data is recall-based, and contains agricultural data from the 2009 summer/fall season to the 2013 spring/summer season. As with our Kenya sample, we cannot be sure of the existence or direction of any recall bias, though trends in our dataset roughly mirror trends in aggregate data for Vietnam.[7]

The Vietnamese dataset is also not balanced, because we have data only on seasons that farmers report remembering; the probability of reported recall is decreasing with time. Furthermore, the number of seasons remembered per farmer is correlated with fertilizer use and probability of biochar adoption (result not shown). To avoid potential bias, we therefore compare estimates using all available data against estimates using season 5-8, which most farmers recalled.

Summary statistics are provided in table 3.1, by biochar adopter/non-adopter, and by season of adoption for adopters. Modest heterogeneity in time-invariant variables is apparent; for instance non-adopters have less land on average than adopters, and there is some heterogeneity in gender and income between adoption groups. More importantly however for the validity of our fixed-effects idetification strategy, we have no evidence of different pre- or post-adoption trends among any of our groups in any of the seasons for which we have data (table 3.3). This suggests that unobserved time-varying factors affecting yield may be balanced across adoption groups, improving the credibility of the common trends assumption we'll require for causal interpretation of our estimates of effects of adoption.

3.2.2 Econometric Models

We assume an underlying standard formulation of the production function $q = f(\mathbf{I}, B)$ in which $\mathbf{I}$ is a vector of inputs such as fertilizers, manure, labor, etc., and B indicates biochar use. We use a panel/longitudinal regression model to assess the impact of biochar adoption on average yields. Given that we do not have data on the actual amounts of the inputs used in our Kenyan data [8], we specify the estimable version of our production function as dependent on fertilizer expenditures F: $yield_{it} = \delta B_{it} + F_{it}'\gamma + \mathbf{X}_{it}'\beta + \epsilon_{it}$.

[7]In 2010, 2011 and 2012, Vietnam produced 4.0, 4.23, and 4.36 $\times 10^7$ tons of rice[1], while average yields in our sample were 8.86, 9.30, and 9.62 tons/hectare. In percentage terms, aggregate yields increased 4.9% and 8.5% from 2010 in 2011 and 2012, while average yields in our sample increased 5.7 and 9.1% over the same period.

[8]Though we lack data on fertilizer amounts used in our Kenyan dataset, we have aggregate data on prices of major fertilizers in the region for each time period in our dataset. We use this aggregate data to normalize expenditures for each time period to units of fertilizer in the final period of our dataset. To maintain comparability with our Kenyan data, we follow a similar approach with our Vietnamese dataset – where we have data on quantities of individual fertilizers used.

The subscripts i and t represent individual farmer i at time period t, B is a variable that identifies biochar adopters, and $\mathbf{X}$ is a matrix of covariates used only for the Vietnamese data, summarized in table 3.4. Employing a standard error components decomposition $\epsilon_{it} = \alpha_i + \rho_t + u_{it}$, we control for farmer-specific intercepts α_i and time effects ρ_t, which allows for causal interpretation of δ if $\mathbb{E}[u|T, \mathbf{X}, \alpha, \rho] = 0$. This is violated if farmers select into adoption based on un-modeled time-varying factors, which is, given our data, not a testable assumption. We provide partial evidence that it is not violated by comparing trends in the unconditional outcome variable (table 3.3), as described above. In the Vietnamese data, we interact time-period variables with province variables, to account for the geographic separation between regions. We specify our treatment variable B as an indicator, equal to zero in absence of adoption and one in presence of adoption but not back to zero if more biochar isn't added in the subsequent season – due to biochar's persistence in soil.

While the fixed-effects regression above gives unbiased estimates of δ conditional on our assumptions, it does not model heterogeneity in response to biochar or fertilizer. However, there is evidence that benefits of fertilizer and biochar are heterogeneous [25, 32], and that benefits to fertilizer are mediated by soil carbon [6, 7]. We therefore also estimate a mixed model in which our vector β includes both a fixed component and an individual-specific random component representing heterogeneity in the response to biochar, fertilizer, and the interaction between the two, mediated by unobserved time-invariant factors that are controlled-for in our fixed effects α_i:

$$yield_{it} = \alpha_i + P_t + (\beta_1 + \beta_{i1})B_{it} + (\beta_2 + \beta_{i2})\text{Fert. Exp.}_{it} + (\beta_3 + \beta_{i3})\left(\text{Fert. Exp.}_{it} \times B_{it}\right) + \epsilon_{it} \tag{3.1}$$

Thus specified, model 3.1 captures not only the average effects of biochar adoption but also the distribution of the impact of biochar adoption. For estimation, we exploit the link between penalized regression and random effects[110, 111], specifying random effects as interactions of a matrix of cross-sectional fixed effects with B, Fertilizer Expenditures and $B\times$ Fertilizer Expenditures, and shrink them via ridge penalties chosen to maximize the restricted maximum likelihood smoothing parameter selection criterion [53].

Finally, we provide robustness checks against potential shortcomings with model 3.1. While we model yields as continuously-distributed, they are clearly non-negative. We show, for our Kenya data, that several approaches towards accounting for this fact yield equivalent results, including the addition of a small constant before taking the log, and specification of discretized yields using negative binomial distribution. This is presented in appendix B.

3.2.3 Optimization

We use estimates of 3.1 to investigate the effect of a carbon price on the profitability of biochar adoption. In our model, farmers choose the proportion of land on which they apply biochar as well as the amount of inorganic fertilizer applied to each plot. We use individuals' maximum observed fertilizer expenditure as a proxy for their budget constraint and assume the cost of a 25kg bag of biochar to be \$4.5 bag, which is the price observed in the region at the time of this study[9]. If we assume (conservatively[10]) that this biochar is half-carbon by mass, and that half of this carbon is sufficiently stable to be considered "sequestered," then each 25kg sack represents 22.9kg CO_2equivalent. This corresponds to 0.45 tons CO_2e/ha, or 183kg CO_2e/acre at recommended application rates.

We simulate a program that subsidizes biochar[11], and represent the optimization problem formally as:

$$\begin{aligned} profit_i = \underset{a, c_F^B, c_F^0}{argmax} \quad & \left(a\left(\hat{f}_i(c_F^B)p - c_B + S\right) + (1-a)\left(\hat{f}_i(c_F^0)p - c_F^0\right)\right) \\ & s.t. : a(c_F^B + c_B - S) + (1-a)c_F^0 \leq C_i \end{aligned} \tag{3.2}$$

Here, a is the proportion of land with biochar, $\hat{f}_i()$ is our fitted model 3.1, c_F^B and x_F^0 are the costs of using fertilizers in presence and absence of biochar, p is the output price, C^B is the price per-hectare of biochar, S is the corresponding subsidy based on a price for

[9] As of this writing – three years after our dataset was collected – a project in the region has begun producing biochar from sugar cane residues for sale to farmers in the region. Recommended application rates are 2 sacks per 1/4 acre, or slightly less than .5 tons/ha.

[10] Biochars are heterogeneous, with larger or smaller carbon fractions depending on feedstock and pyrolysis technology. Biochar stability is difficult to estimate given its long residence time in soil, the lack of good natural experiments spanning centuries, and the complexity of physical, chemical, and biological transformations and translocations in soil. However, current models of biochar decomposition generally divide it into "labile" and "stable" pools based on the heterogeneity of carbon compounds comprising the material. Decomposition rates of the labile pool are generally of the same order of magnitude as other organic substances in soil, while that of the stable pool is often several orders of magnitude less. Mean residence times for this pool are generally estimated to be on the order of centuries to millennia[38, 112].

[11] Alternatively, we could have considered a program which monitored behavior with regard to biochar application to soil. However, the cost of measuring compliance among many small-scale farmers would be high, and likely counterproductive given that these monitoring costs would reduce the proportion of funds available to actually defray the cost of biochar. Second, even at a very high carbon price of 100USD/ton, a farmer applying biochar to a hectare of land would receive a subsidy of 45.2USD, which is less than the price of an hectare's-worth of biochar (90.9USD). Thus there are no realistic scenarios in which adoption of biochar would lead to a cash payout, rather than a simple defrayment of costs. Third, a program that monitors input usage and provides payouts based on them could value the CO_2e implicit in any reductions in fertilizer use attributable to substitution with biochar. Any such program would face serious attribution issues, potentially create perverse incentives for farmers to stop using fertilizer in order to claim a payout, and also conflict with the existing policy priority of expanding use of inorganic fertilizer in sub-Saharan Africa as a means of improving food security. Given these factors, we consider input subsidies – rather than behavior-based cash payouts – to be a more realistic model for how a carbon price might influence a market for biochar.

carbon, and C_i is a farmer i's budget constraint. Input and output prices are summarized in table 3.5. Optimization is performed by a simple gridsearch across the feasible region of the parameters, and p is specified in 1/8-acre increments – reflecting 25kg sacks of biochar.

The objective function depends critically on $\hat{f}_i()$, our fitted model 3.1. To characterize uncertainty in this estimate, we adopt a Bayesian view of its estimated parameters – treating them as being distributed multivariate-normally, with their mean defined by $\hat{\beta}$ and their covariance defined by our estimated (cluster-robust) variance-covariance matrix[12][13]. We then take 1000 draws from this distribution, conducting optimizations for each draw of $\hat{f}_i()$ as the emprical distribution of the optimized parameters across the 1000 draws of β. We do so for a gradient of carbon prices, from \$0-\$100/ton, and calculate confidence/credible intervals using the quantiles of the predicted optima defined by different draws of β. Our estimands of interest are, as a function of carbon price, (1) the proportion of farmers for whom it is optimal, at varying levels of confidence, to adopt biochar in any quantity, (2) change in expected profits, (3) the proportion of land that receives biochar in the first year of the program.

3.3 Results

Estimates for the Kenya and Vietnam datasets are given in tables 3.6 and 3.7. For the Kenya sample, biochar adoption is associated with a .156 ton/ha increase in maize yields, as well as a \$26.2/ha decrease in fertilizer expenditure. When holding fertilizer expenditure constant and also controlling for the interaction with fertilizer (which is negative and marginally statistically significant), the estimate increases to .211 ton/ha. With pre-adoption yields averaging 1.3 tons/ha, the estimate of the average treatment effect for biochar-adopting farmers is approximately 11%. In our Vietnam sample, we do not detect an effect of biochar adoption on yields, though we do detect a decrease in fertilizer expenditure in regressions run only on season 5-8. The effect on yield remains statistically insignificant after controlling for available time-varing variables.

[12] A frequentist interpretation of this procedure would be that we are drawing from the estimated distribution of parameter estimates that we'd obtain after drawing many different samples from the meta-population from which our data was drawn.

[13] We estimate a cluster-robust variance-covariance matrix for models containing penalized terms as $\hat{\mathbf{V}}_{\hat{\beta}} = \left(\mathbf{X}'\mathbf{X} + \hat{\lambda}\mathbf{S}\right)^{-1} \sum_j \mathbf{X}_j' \hat{\epsilon}_j \hat{\epsilon}_j' \mathbf{X} \left(\mathbf{X}'\mathbf{X} + \hat{\lambda}\mathbf{S}\right)^{-1}$, where $\hat{\lambda}$ is the estimated term-wise smoothing paramter and $\mathbf{S}$ is a penalty matrix – controlling which components of $\mathbf{X}$ are penalized. Heteroskedasticity-robust F-tests on purely parametric model components are then based off of the above matrix. Tests of term collections containing penalized elements are done by replacing our cluster-robust estimate of $\mathbf{V}_{\beta}$ with the standard one in calculating Wood's [80] reduced-rank Wald statistic $T_r = \hat{f}_j \hat{\mathbf{V}}_{f_j}^{r-} \hat{f}_j$ where $\hat{\mathbf{V}}_{f_j}^{r-}$ is a rank-r pseudoinverse of $\hat{\mathbf{V}}_{f_j}$. Doing so accounts for the fact that effective degrees of freedom for a penalized term or a term containing penalized components are fewer than the dimensions comprising the portion of the variance-covariance matrix corresponding to the terms for which the test is desired.

Mixed model estimates of fixed terms are generally consistent with OLS estimates, with slightly lower (higher) estimates of the average effects of biochar (fertilizer). In both the Kenya and Vietnam cases, the random effects are jointly statistically significant. Their distributions and correlations are given in figure 3.2. In both cases, there is sizable heterogeneity in response to biochar, which is negatively correlated with heterogeneity in response to fertilizer. There is little heterogeneity in the interaction between biochar and fertilizer, though it is positively correlated both with response to biochar and fertilizer.

Profit maximization based on simulations of our production function suggests that biochar has a relatively small effect on profits, and a small negative effect on fertilizer expenditure (figure 3.6). Notably, even in presence of compensation for carbon sequestration, results indicate a low sensitivity of profitability to carbon price (figure 3.3). We are more confident than not that biochar would be profitable to adopt for 23% of our sample, given no carbon subsidy. Conditional on profitability however, most farmers in most scenarios apply biochar to only small portions of their land (3.4), irrespective of carbon price. This proportion increases to 47% with a subsidy of \$100/ton CO_2e, though neither proportion is statistically different from zero at a confidence level of 95%. Profits and fertilizer expenditure appear to be relatively insensitive to carbon price (figure 3.5). In particular, the benefit of a subsidy is for most farmers a small fraction of the cost of providing it. We do not conduct optimizations for the Vietnam sample, as average effect on yields is not statistically significant, and heterogeneous response to biochar and its interaction with fertilizer are inestimable for those in our sample who did not receive it.

3.4 Discussion

Biochar has elicited substantial interest as a means of sequestering carbon while improving agricultural yields. Translating technical potential into impact requires wide-scale adoption, and wide-scale adoption will generally be predicated on economic feasibility. A priori, we were hopeful to find evidence for such feasibility, given that biochar is one of the approaches to climate change mitigation expected to generate positive economic returns in the short run. In particular, in Kenya we were intrigued by the possibility that biochar might help intensify low-input, low-yielding agriculture, while improving yield responsiveness to inorganic fertilizer. In the more input-intensive Vietnamese agricultural system, we were interested in the possibility that biochar might maintain high yields while reducing the use of inorganic fertilizer with gains in efficiency and a reduction in GHG emissions.

Evidence in support of these potentials is mixed. We find that biochar improves crop yields in western Kenya, but find little evidence for complementarity with fertilizer. Even though our findings could be contradicted by the dynamics of biochar's influence on agro-ecosystems over longer periods, we cannot conclude that biochar provides a mean of stimulating agricultural intensification in the short run. We do find evidence that response to biochar is negatively correlated with response to fertilizer – echoing findings by

Marenya and Barrett[6, 7]). In Vietnam case, we find no evidence that biochar increased average yields, though substantial heterogeneity in the random coefficient suggests that there may be circumstances in which biochar applications are valuable; further research is needed. In Kenya, we find that average yields increased while average fertilizer expenditure decreased. If biochar were available freely, this would imply an increase in agricultural profit. However, it is unlikely that biochar would be available at no cost at wide scale to farmers, and estimated optimal input mixes – subject to budget constraints – suggest that its provision would change profits little, with or without a subsidy. And while according to our results, one quarter to one half of farmers may find it profitable to adopt (depending on compensation for carbon sequestration), this adoption might involve substitution with inorganic fertilizer, which also improves yields but doesn't appear to have a statistically significant complementarity with biochar. In focusing on optimality, we abstract from more complex issues related to technology adoption – of which profitability is only one. Additional work and data is needed to determine how closely our modeled input decisions capture farmers' decision-making processes and what sorts of interventions might be required to stimulate adoption. We are aware that our results are likely to be sensitive to maize and fertilizer prices. However, while changes in these factors may change the magnitudes of our estimates, we don't expect that they will substantially change results qualitatively.

Our results are weakened by omitted variables that vary with time. We lack data on seed variety, labor, other crops, or other inputs like tractors or animal rental. Omission of these variables may bias our estimates, to the extent that they are correlated both with yields and the timing of biochar or fertilizer use. Hybrid seeds tend to have higher response to fertilizer than traditional varieties. If farmers used more of these due to any wealth effect stemming from previous-period biochar adoption, our estimates might overstate the effect of fertilizer. Heavily-fertilized fields require more weeding, which might require more labor, which might be out of reach for budget-constrained households. This would generally bias the effect of fertilizer use upwards. Also, if wealth effect from biochar adoption led to farmers increasing the amount of land planted, then we'd expect the estimate of biochar's effect to be biased downwards, given the prevalent inverse relationship between yields and farm size in African agriculture. Finally, we represent biochar use as a dummy variable, rather than a continuous variable applied heterogeneously to different plots within a farm. Any bias inherent in the unobserved proportion of their land to which biochar was applied will be downward, as we can't distinguish between plots with and without biochar. Altogether, we don't know how serious the effect of omitting these variables may be.

Several related points beyond the scope of our empirical analysis are worth noting for any projects aiming to promote biochar utilization. First, biochar's viability is conditional on ecological constraints. In particular, biochar requires adequate biomass to serve as feedstock, a constraint that appears to be met in our contexts[113], but which may not be in all places, even where biochar may be agronomically beneficial. Indeed, removal of crop residues from fields – against a counterfactual of in-field decomposition – implies an export of plant nutrients and an increased erosion risk, even if future external nutrient additions may be more efficiently used because of any influence of biochar. These constraints will vary with soil type, agro-climatic zone, prevalent cropping systems, and other factors. Biochar's viability will thus by highly context specific.

Second, even where feedstock is available, biochar production requires some technology for pyrolysis. These can be centralized or farm-scale. Centralized systems will generally be costlier, more efficient, and require substantial transportation of bulky feedstocks to and from farms – though they may capture useful energy in the process[36, 92, 114–116]. In the absence of strong market demand by farmers and cost-effective and reliable systems for moving large quantities of biomass, the relative advantage of using feedstocks to produce biochar rather than electricity remains unclear. Farm-scale systems include biochar-producing stoves[117] and larger TLUD-kilns[118]. Biochar-producing stoves can produce useful energy for cooking while remediating problems with indoor air pollution common in developing country contexts[119]. However it is not clear that the amounts of biochar produced will be agronomically relevant. Moreover, evidence on the adoption and sustained use of improved cookstoves in general is mixed[120–123], and it is not clear that pyrolytic stoves are substantially better than other sorts of improved cookstoves[119].

Third, feedstock availability and labor are commonly seasonal in agriculture, in ways that may work against the practicality of producing large quantities of biochar. Crop residues are generally available once or twice per year, which requires that biochar producers either have substantial capacity for rapid production, or capacity to store large quantities of biomass while keeping it dry. Seasonal inactivity and/or material storage or drying costs are likely to impose constraints on centralized systems, due to lower capacity factors and marginal costs, respectively. Timing of harvests and seasonal wet and dry periods may shrink the window in which material can be pyrolyzed in some regions; in Kenya, the short rains follows quickly after the long rains harvest – wetting the material from which biochar could be produced. Small-scale systems – such as TLUD kilns – are fairly labor intensive, and are likely to be constrained where agricultural wages are high. These constraints can be overcome somewhat by decentralized production, though decentralized production would be more likely to face behavioral technology adoption constraints.

Given these constraints and given the relative invariance of optimal biochar usage to price (within a realistic range), our results suggest that any approaches to promoting biochar may be more successful by focusing on removal of barriers to biochar provision in agricultural markets in the region, in addition to simple subsidization. These factors would clearly vary by context, but might include extension or other information campaigns, support for nascent biochar producing businesses or nonprofits, and identification of regions where the above-discussed constraints don't bind as severely.

3.5 Conclusion

We find that biochar increased yields in western Kenya, and that its profitability may be moderately sensitive to a carbon price. However, we don't find evidence that biochar improves response to fertilizer, and as a consequence we find little evidence that its adoption would substantially change agricultural profit, given substitution with little complementarity. In Vietnam, we find no average increase in yields, though in both contexts, we find substantial heterogeneity in response. This heterogeneity invites further research; given that biochar appears to work in some contexts more than others, identification of its niche may maximize the benefits of any programs seeking to leverage it for agricultural development and climate change mitigation.

3.6 Figures and tables

Table 3.1: Summary statistics for the Vietnam sample, disaggregated by season of initial adoption.

		Yield		FertExp		Gender		Age		HouseholdSize		LandRice		LandTotal		NonFarmIncome	
AdoptionSeason	n	Mean	Sd	Mean	Sd	Mean	Sd	Mean	Sd	Mean	Sd	Mean	Sd	Mean	Sd	Mean	Sd
Fall 2009	9	4.642	0.7762	192.5	40.95	0.0000	0.0000	39.78	9.244	4.333	1.323	0.2240	0.10147	0.6791	1.0754	2681	1548
Spring 2010	31	5.319	0.8763	231.7	69.22	0.2903	0.4614	48.35	9.911	5.000	1.317	0.2001	0.11472	0.3171	0.2992	4763	2840
Fall 2010	10	4.827	1.2530	205.0	71.95	0.2000	0.4216	46.90	10.999	4.300	1.567	0.2146	0.09182	0.2776	0.1088	3908	2995
Spring 2011	48	5.467	0.8422	214.9	79.76	0.5625	0.5013	51.19	10.776	4.896	1.477	0.2036	0.09584	0.3073	0.3419	3917	2263
Fall 2011	29	5.533	1.1337	243.6	93.91	0.5172	0.5085	51.96	11.328	4.241	1.300	0.2003	0.09305	0.2559	0.1481	3522	1669
Spring 2012	47	5.133	1.1628	221.2	83.75	0.2766	0.4522	47.11	11.469	4.702	1.350	0.1790	0.08120	0.2662	0.3281	3549	1868
Fall 2012	39	4.979	1.0351	186.4	67.73	0.2308	0.4268	44.59	10.728	4.590	1.117	0.2430	0.19556	0.6253	1.1570	4040	2883
Spring 2013	8	5.116	1.3637	188.3	115.84	0.0000	0.0000	43.50	10.784	4.750	1.389	0.2205	0.14584	1.7525	3.4245	2735	1507
Non-adopters	75	4.941	0.8725	283.7	88.84	0.3333	0.4746	51.61	11.593	4.333	1.510	0.2142	0.10909	0.2401	0.1341	4208	2458
All	296	5.151	1.0107	232.5	87.65	0.3378	0.4738	48.85	11.339	4.588	1.392	0.2083	0.11806	0.3710	0.7781	3929	2376

Table 3.2: Summary statistics for the Kenya sample, disaggregated by season of initial adoption.

		Yield		FertExp		Gender		Age		HouseholdSize		LandMaize		LandTotal		NonFarmIncome	
AdoptionSeason	n	Mean	Sd	Mean	Sd	Mean	Sd	Mean	Sd	Mean	Sd	Mean	Sd	Mean	Sd	Mean	Sd
Long Rains 2008	8	0.9769	0.2786	19.47	7.553	0.5000	0.5345	44.50	15.94	7.500	5.606	0.6323	0.4978	3.781	2.901	811.8	1188.3
Short Rains 2008	2	1.6680	1.1008	55.24	53.447	0.5000	0.7071	25.00	*NA*	5.000	1.414	0.4047	0.0000	2.500	2.121	847.1	1197.9
Long Rains 2009	124	1.7195	1.0630	74.66	62.821	0.3548	0.4804	44.01	11.96	6.073	1.976	0.4202	0.3862	2.147	1.840	936.3	1144.3
Short Rains 2009	36	1.5005	0.8445	53.19	54.967	0.3333	0.4781	44.39	11.61	6.500	2.678	0.4497	0.2665	2.507	1.613	1409.7	2230.5
Long Rains 2010	125	1.7081	1.1928	65.05	64.849	0.4677	0.5010	41.70	11.19	6.384	2.436	0.4678	0.3404	2.750	3.327	1524.2	2357.6
Short Rains 2010	7	0.9452	1.2337	44.58	86.134	0.4286	0.5345	40.00	15.51	5.286	2.059	0.4914	0.3849	3.357	3.913	474.5	462.9
All	302	1.6630	1.0973	66.87	63.179	0.4053	0.4918	42.95	11.81	6.265	2.408	0.4506	0.3565	2.513	2.627	1231.3	1896.8

Table 3.3: Tests for un-common trends in outcome variables, pre- and post-adoption, for our Kenya and Vietnam samples. Tables give p-values of F-statistics associated with regressions of the form $trend_{it} = yield_{it} - yield_{it\pm1} = \mathbf{A}'\beta + \epsilon$, where $\mathbf{A}$ are factor variables for the season of initial adoption. Each regression omits data from groups who adopted in the current or previous season. Overall, inter-seasonal changes in yield are not staistically distinguishable between groups defined by season of adoption.

	Season						
	1-2	2-3	3-4	4-5	5-6	6-7	7-8
Kenya	0.866	0.606	0.428	0.151	0.488		
Vietnam		0.314	0.358	0.304	0.888	0.668	0.700

Table 3.4: Covariates available for the Vietnamese data.

	Mean	SD
Fertilizer Expenditure (USD)	175.7173	138.0754
Irrigation Expenditure (USD)	10.4492	5.9096
Labor Expenditure (USD)	193.5237	82.1981
Other Expenditure (USD)	79.0390	55.7009
Manure (kg/ha)	867.8337	1633.8788
Area farmed (ha)	0.1987	0.1034

TABLE 3.5: Fertilizer and grain prices in nominal USD during the periods spanned by our dataset. Per-nutrient prices calculated assuming DAP rating of 18-46-0, NPK rating of 15-15-7.5, urea rating of 46-0-0, phosphate as single-superphosphate 0-18-0, and potassium as 0-0-60.

Kenya						
	2008		2009		2010	
	Long Rains	Short Rains	Long Rains	Short Rains	Long Rains	Short Rains
Sale price of maize (kg)	0.26	0.16	0.37	0.26	0.42	0.26
Price of DAP fertilizer (kg)	0.61	0.56	0.82	0.75	1.13	1.06
Price per kg N	*3.40*	*3.14*	*4.58*	*4.18*	*6.27*	*5.88*
Price per kg P_2O_5	*1.33*	*1.23*	*1.79*	*1.64*	*2.46*	*2.30*
Vietnam						
	2009	2010	2011	2012	2013	
Sale price of rice (kg)	0.24	0.25	0.25	0.25	0.28	
Purchase price of NPK Fertilizer (kg)	0.21	0.14	0.18	0.19	0.24	
Price per kg N	*1.43*	*0.95*	*1.21*	*1.27*	*1.59*	
Price per kg P_2O_5	*1.43*	*0.95*	*1.21*	*1.27*	*1.59*	
Price per kg K_2O	*2.86*	*1.89*	*2.41*	*2.54*	*3.17*	
Purchase price of Urea (kg)	0.33	0.32	0.50	0.55	0.61	
Price per kg N	*0.72*	*0.70*	*1.09*	*1.19*	*1.32*	
Purchase price of Phosphate fertilizer (kg)	0.19	0.11	0.19	0.19	0.18	
Price per kg P_2O_5	*0.41*	*0.24*	*0.41*	*0.41*	*0.39*	
Purchase price of Potassium fertilizer (kg)	0.48	0.45	0.58	0.58	0.62	
Price per kg K_2O	*0.79*	*0.75*	*0.97*	*0.97*	*1.03*	

TABLE 3.6: Empirical results from Kenya sample. Standard errors adjusted to account for cross-sectional autocorrelation.

	Dependent variable:			
	Yield FE-OLS	Fert. Exp. FE-OLS	Yield FE-OLS	Yield FE-LMM
	(1)	(2)	(3)	(4)
Biochar	0.156***	−26.218***	0.211***	0.171**
	(0.060)	(6.413)	(0.077)	(0.071)
Fert. Exp.			0.002**	0.004***
			(0.001)	(0.0005)
Biochar x Fert. Exp/			0.00003	0.0003
			(0.001)	(0.001)
Observations	1,295	1,270	1,267	1,267
Adjusted R^2	0.666	0.562	0.684	0.768
AIC			2506	2203
Note:		*p<0.1; **p<0.05; ***p<0.01		

TABLE 3.7: Empirical results from Vietnam sample. Standard errors adjusted to account for cross-sectional autocorrelation. Right panel excludes data from seasons 1-4.

	Dependent variable:						
	Yield FE-OLS	Fert. Exp FE-OLS	Yield FE-OLS	Yield FE-LMM	Yield FE-OLS	Fert. Exp FE-OLS	Yield FE-OLS
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
BC	0.041	−2.991	−0.004	−0.033	−0.109	−10.056*	−0.029
	(0.099)	(4.715)	(0.290)	(0.223)	(0.147)	(5.697)	(0.443)
Fert. Exp.			0.001	0.001*			0.001
			(0.001)	(0.001)			(0.002)
Other. Exp.		0.039	−0.0003	−0.0002		−0.013	−0.001
		(0.042)	(0.001)	(0.001)		(0.064)	(0.001)
Irrigation Exp.		0.154*	0.001	0.0003		0.190*	0.003
		(0.092)	(0.002)	(0.002)		(0.114)	(0.003)
Manure (ton)		−0.016	0.040*	0.041*		0.525	0.023
		(1.049)	(0.022)	(0.021)		(1.287)	(0.029)
Labor Exp.		0.047*	0.001**	0.001**		0.038	0.001
		(0.027)	(0.0004)	(0.0004)		(0.031)	(0.001)
Land in Rice (ha)		62.677	−3.849***	−4.062***		68.307	−4.454***
		(66.504)	(1.234)	(1.233)		(78.337)	(1.515)
BC x Fert. Exp			0.0002	0.0004			−0.0003
			(0.001)	(0.001)			(0.002)
Observations	1,539	1,541	1,537	1,537	1,093	1,094	1,092
R^2	0.667	0.793	0.678		0.697	0.846	0.708
Adjusted R^2	0.578	0.736	0.590	0.659	0.578	0.784	0.589
AIC			3794.14	3587.60			

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Figure 3.1: Estimates of parametric terms (plots 1 and 2), distributions of random coefficients (plots 3-5)

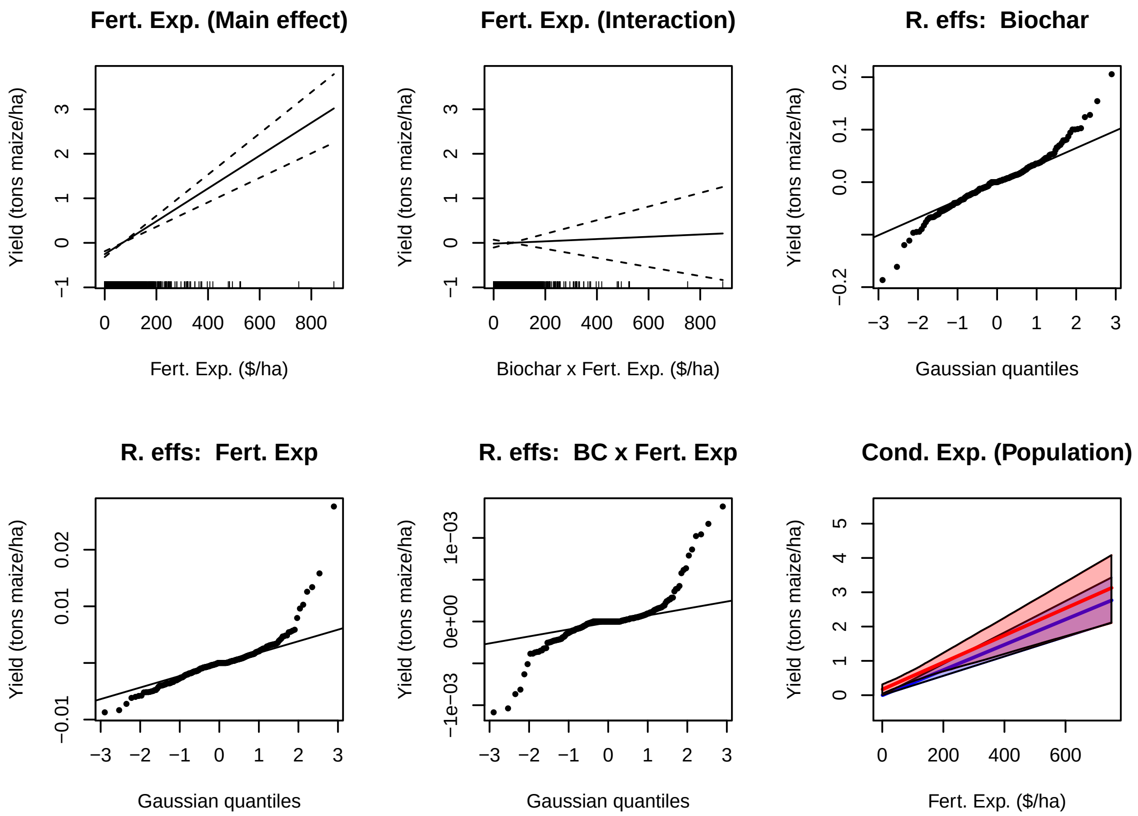


FIGURE 3.2: Correlations between random effects predictions by individual. Linear fits control for the excluded random effect (i.e.: $Biochar \times Fert.Exp$ where biochar effects are plotted against fertilizer effects), and set the excluded to zero. Note differing scales.

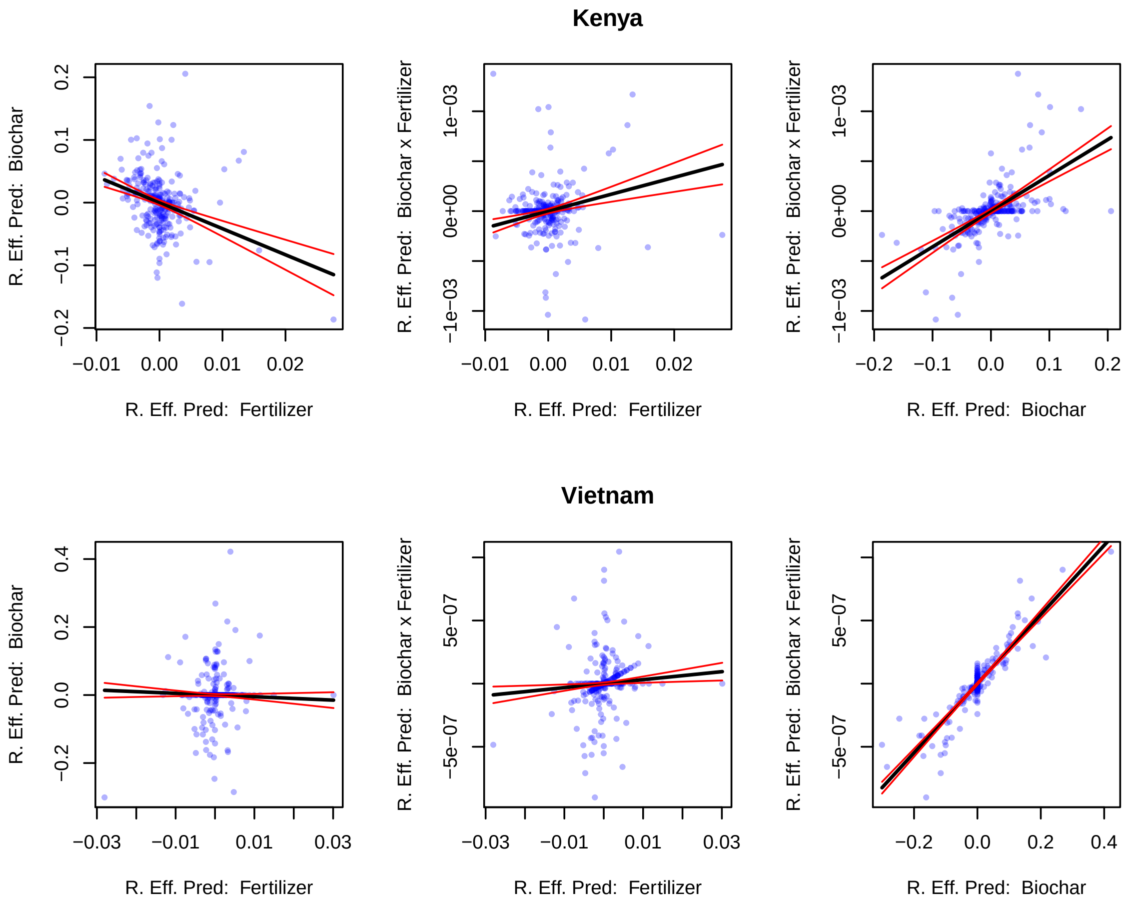


Figure 3.3: Left: proportion of our Kenya sample for who biochar adoption would be optimal from a profit-maximizing perspective, over at least 1/8 of their farm, as a function of carbon price. The red line give the proportion for whom we are 95% confident of optimality, the blue line 75%, and the green line gives the proportion for whom we are more confident than not. Right: assuming profit maximization, the proportion of land receiving biochar as a function of carbon price. Solid line gives median estimate; dotted lines give 95% confidence intervals.

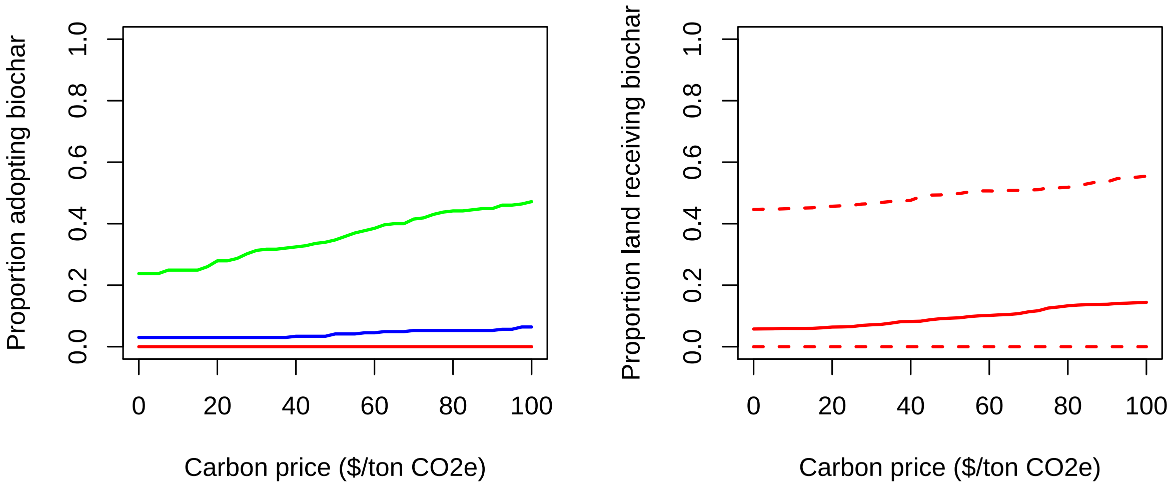


FIGURE 3.4: Conditional on biochar profitability, the estimated optimal proportions of land to which farmers in our sample would apply biochar, at no subsidy and at \$100/ton CO2e.

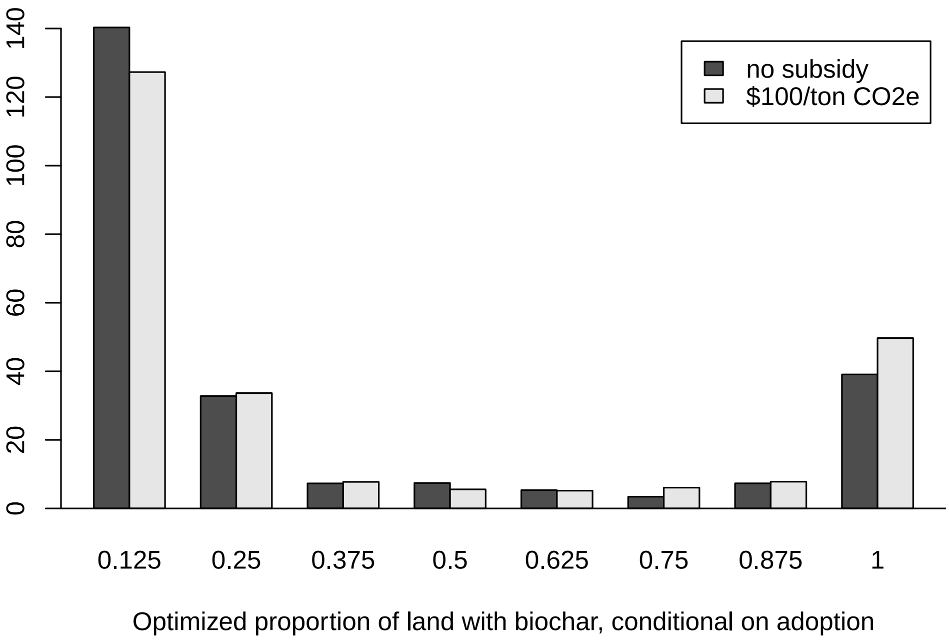


Figure 3.5: Distributions of average change in optimized agricultural profit and fertilizer use as functions of carbon price.

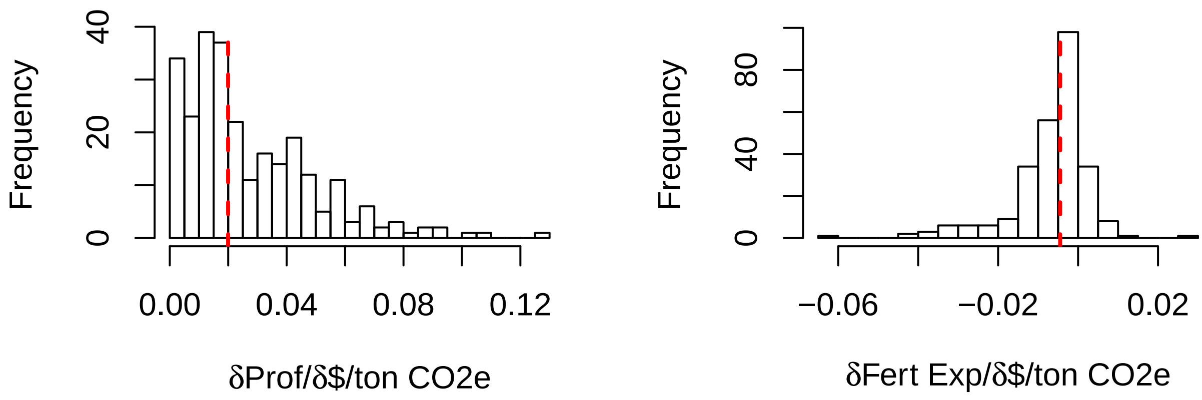
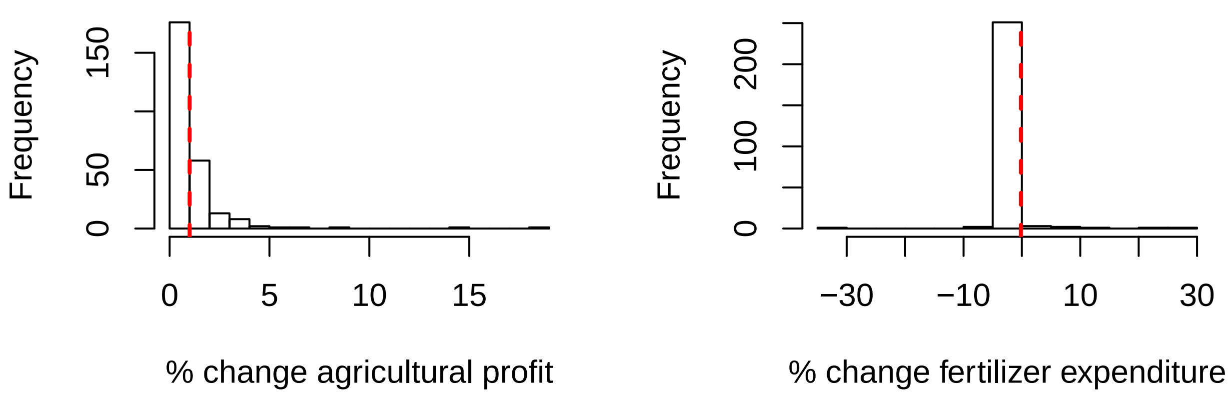


Figure 3.6: Distributions of the average difference in optimized profits and fertilizer expenditures between scenarios where biochar is available with no subsidy, and scenarios in which biochar is unavailable. Dotted red lines give means.



Chapter 4

Subsidies, Social Learning, and Charcoal Dust: Impact and Adoption Dynamics of Biochar in Western Kenya

Abstract
I report the results of a Kenyan field experiment on adoption and impact of biochar, an agricultural input that improves fertility in poor soils while sequestering carbon. I randomly assigned prices, demonstrations, and risk-free trials. Yields increased by 37% and 50% in the first two seasons, and response to inorganic fertilizer improved. However, uptake was 2.6% and 10% respectively. Farmers were highly price-sensitive. Social network effects were marginally significant, but positive at low penetration and negative at high penetration. Given uptake well below the social optimum, subsidies for biochar appear justified from a social cost/benefit standpoint.

4.1 Introduction

Since independence, African agricultural productivity and input use has grown very little, while the population remains overwhelmingly rural and poor[17]. This stagnation is particularly stark in contrast to the massive productivity gains experienced by most of the rest of the world during the same period. Driven by development of hybrid seed and inorganic fertilizer, the post-WWII "green revolution" is credited with much of the gain experienced elsewhere, and considerable attention has been paid to whether and how this experience might be replicated in sub-Saharan Africa[124–127].

Fertilizer's profitability [25] as well as its agronomic efficacy [12, 128] can be quite variable in the region. This profitability is often increasing in by soil carbon content, particularly in the geologically weathered soils common to much of sub-Saharan Africa [6, 7, 27]. Increasingly, institutional and academic discourses around agricultural development in the region have focused on "climate smart" agriculture – a term encompassing the integration of climate change mitigation and adaptation into development assistance in the sector[129]. Climate smart agriculture rests somewhat uneasily with the aim of agricultural development through increased use of inorganic fertilizer, as inorganic fertilizers can be relatively greenhouse gas-intensive in both production and use. However, measures that increase soil carbon content can mitigate climate change via carbon sequestration, while potentially stimulating agricultural intensification by improving responsiveness to fertilizer and thereby the efficiency of its use.

The past several decades have seen substantial agroecological research around this prospect. Approaches have included agroforestry, conservation agriculture, composting/manuring, among several others[81, 130, 131]. Increases in yields are commonly strong, but benefits can take several years to materialize, and adoption rates are often low while disadoption rates among early adopters are high[132, 133]. Particularly in the context of growing opportunities to finance such projects under both voluntary carbon markets and potentially emerging international climate change mitigation agreements [134, 135], evidence is needed on how to use these resources most efficiently to stimulate dissemination, yet relatively little literature focuses on this area[136].

Concomitantly, problems of technology adoption in African agriculture have received substantial attention in development economics – largely in isolation from work in climate-smart agriculture. Following the liberalization of agricultural input markets in the 1980s and 90s, much of this work has focused on means of stimulating adoption of improved agricultural technologies while avoiding the use of blanket subsidies, which were considered expensive, inefficient, and potentially regressive in the way they were implemented[2]. Given that the technologies subsidized were generally profitable in expectation at market price, the goal has been to identify what barriers impede adoption, and how those might be overcome[3]. These barriers vary in scale from those mediated by national policy (land and labor market factors, infrastructure, etc.), to those that are more microeconomic in nature, such as information, credit availability, risk attitudes, and other behavioral factors. A large and growing body of research – often using experimental methods – has found that relaxing these micro-scale barriers has substantially stimulated adoption, while elucidating mechanisms by which they have worked.

This paper focuses on the intersection of these two issues. I provide the first rigorous evaluation of the agronomic efficacy and economic utility of biochar[35] – a soil amendment that can increase yields, persistently improve soil fertility, and sequester carbon – in smallholder agriculture. This is distinct from the controlled agronomic studies that have comprised the literature to date. While such studies are well-suited to advancing mechanistic understanding of biochar's activity in agroecosystems, they are often conducted in controlled settings that bear little resemblance to the contexts in which smallholder farmers operate. In particular, I seek to determine whether the yield impacts seen in agronomic studies are replicated on the fields of farmers, and whether biochar's benefits

– in terms of yields, soil fertility management, and carbon sequestration – are sufficiently large to justify the cost of its production.

Agronomic benefit and economic viability do not guarantee rapid adoption however[3]. Conditional on these two factors, rates of takeup may still be quite slow. Farmers may lack information about the new technology – how it works, what mediates its benefits, the extent of its benefits, or even its existence. In other circumstances, technologies that are beneficial long-term investments may not meet immediate livelihood or consumption needs. Particularly where technologies have beneficial environmental externalities, policymakers may wish to speed rates of dissemination. Yet it is often unclear how to do this efficiently and effectively. This paper seeks to contrast three methods that have been researched independently, and implement them in the specific context of biochar – simple price subsidies, stimulation of social and experiential learning via the provision of demonstration plots, and risk-free trial offers whereby farmers receive biochar upfront, and can choose to pay for it after harvest if they find that it was valuable and wish to buy more in the future.

I find that yields on farmers fields are consistent with, or greater than, results from agronomic studies, and that biochar improves response to fertilizer – even at the very low application rates used in this study – about 0.5 tons per hectare, which I found to be approximately the amount that can be made from the crop residues of a typical maize field in the region. My data suggest that biochar is highly profitable for the median farmer to adopt, at a price similar to the cost of its production, even ignoring any benefits stemming from carbon sequestration and fertility benefits that last beyond the first year.

In spite of this however, adoption rates were quite low. While farmers were highly responsive to risk-free trial offers, most did not repay them. Adoption was much higher in the second season, but was contingent on heavy subsidies. The effect of the social network is zero in the first season and unclear in the second – while point estimates suggest that an increasing share (or proportion) of the network with biochar stimulates adoption at low levels of penetration, point estimates also suggest declining propensity to adopt biochar when more than about 12% of the network has biochar. Estimates are noisy however, and I'm not able to distinguish these estimates from no effect whatsoever at standard levels of statistical confidence.

Given that biochar is unambiguously profitable yet underadopted, low rates of adoption are socially suboptimal. This suggests a role for policy to stimulate adoption. I find that a combination of heavy subsidies and measures to stimulate social learning maximize social return on investment (though I can't say at high confidence that the optimal program would involve no social learning component), but that the amount of carbon sequestered by biochar is insufficient to justify the cost of programs to stimulate its uptake based on its carbon benefit alone. However, while public expenditures to stimulate biochar adoption should yield substantial social return on investment – given its direct agronomic benefit, its complementarity with fertilizer, and its carbon sequestration potential – the large subsidies needed to induce farmers to try a new technology may render biochar a less attractive public investment than simple fertilizer subsidies, in the short term. Biochar may be a more favorable investment when considered over the longer term, when carbon sequestration is valued, or in particularly degraded and/or non-fertilizer-responsive soils. Further research is needed to place these results into an even more dynamical context, given that willingness to pay may change over time.

After introducing biochar (section 4.1.1), I provide a brief literature review of the development economics literature in this space (section 4.1.2). I then describe the context and experimental design (section 4.2), including a brief treatment of the semiparametric statistical methods that I use (section 4.2.2) throughout the paper. The dataset is described in section4.3. I then provide results (section 4.4) and conclude with a brief discussion (section 4.5) focused on implications for policy and for further research.

4.1.1 Biochar

Biochar is the product of thermal decomposition of biomass in the absence of oxygen[35]. Termed "pyrolysis," the process of its production can be given by the following heuristic chemical reaction:

$$CH_2O + \text{Heat} \rightarrow \textbf{Biochar} + H_2 + CH_4 + \text{Heavier hydrocarbons}$$

In essence, heat drives flammable gases and vapor-phase hydrocarbons[1] from biomass (approximated as CH_2O), while the solid that remains is chemically altered, forming black carbon.[2] The key distinction between this process and the process of simple combustion is the absence of ambient oxygen. Absent oxygen, a substantial portion of the feedstock remains in solid form, while heat alters the chemical structure of the remaining carbon atoms into increasingly condensed aromatic structures[3] with increasing temperature. There is no formal distinction between biochar and charcoal, either chemically or in their manner of production. Rather the two are differentiated by their intended end-use, though many forms of charred biomass – made, for example, from leaves, grass, or crop

[1] In traditional charcoal production, these byproducts are vented to the atmosphere, releasing harmful local air pollutants and greenhouse gazes. In more-advanced biochar production systems (which this project used, and which are also more efficient), these flammable gases are combusted to further drive the reaction.

[2] Detailed treatments of this process – of which the production of fuel charcoal is a special case – can be found in Brown [115] and in Antal & Gronli [137].

[3] "Aromatic" carbon refers to carbon that has formed a hexagonal ring of 6 carbon atoms. "Condensed" aromatic carbon refers to collections of ring structures that have fused to one another. Graphene occurs when aromatic carbon is perfectly condensed into sheets of attached rings, while graphite is formed when these sheets layer onto one another. The degree of aromaticity (i.e.: the proportion of carbon in aromatic clusters), and the degree of condensation (i.e.: the proportion of ring vertices that are bonded onto other aromatic rings) varies with a number of parameters of the biochar production process, especially the temperature of production, which typically varies between 300°C and 750°C. Biochar is substantially more aromatic that living plant carbon, but substantially less so than graphene. See the following for more on biochar's chemistry: [58, 76, 138, 139].

residues – would tend to be less useful as a fuel than would char made from large pieces of hardwood, which is more typically used for fuel charcoal.

Like other forms of nonliving organic carbon, biochar will "mineralize" over time, reacting with oxygen to form carbon dioxide. However, biochar does so much more slowly than the dead plant matter from which it is formed[28, 38, 83] – decomposing over timescales ranging from centuries to millennia. Early interest in biochar was catalyzed by the discovery of surprisingly fertile, dark-colored soils in the Brazilian Amazon, which were surrounded by highly weathered and much less-fertile soils[140]. These soils, which are coincident with pre-Columbian population centers, contain carbon dated to several centuries before European contact. Their fertility is attributed largely to addition of biochar, likely via produced by smoldering combustion during the process of land-clearing for agriculture [39, 141, 142].

By slowing the reversion of plant carbon back to the carbon dioxide from which it was formed, biochar suggests itself as a means of climate change mitigation via soil carbon sequestration. Woolf et al.[36] estimate that biochar may have the technical potential to mitigate as much as 10% of current global greenhouse gas emissions anually. Of course, the gap between technical potential and realized potential will depend on deployment and use of biochar production technology, which will in turn depend on economic feasibility. Biochar's costs are generally comprised of (1) the opportunity cost of the feedstock biomass, (2) capital costs of the equipment used to perform pyrolysis, (3) labor costs required to perform pyrolysis, and (4) transportation costs inherent in moving biomass to the equipment and biochar from the equipment. These costs vary substantially by context and by the type of technology used.

On the other hand, biochar's economic benefits are largely a function of its ability to improve crop yields. Meta-analysis of agronomic studies shows that average yields increases range from 5% to 30% over controls, with relatively higher increases in the weathered soils typical of much of the tropics, and lower to negative yield increase in the richer soils that make up much of the world's more important food-producing areas in the temperate zone[32, 44, 45]. Similar to other forms of soil organic matter (SOM), biochar can improve soil fertility across a number of dimensions – improving water holding capacity and soil structure, while reducing acidity and improving fertilizer-use efficiency [37, 47].[4] Importantly, these results have been shown to manifest immediately after application, and also to persist over multiple seasons beyond the season of initial application, offering the prospect of durably increasing soil fertility.

[4] While the sole meta-regression analysis of plant response to biochar did not find evidence that biochar improved plant response to fertilizer[32], that study did not explore potential heterogeneity in mediators of response, and it is possible that it may do so in some contexts and not others. I will test for evidence that it does so in section 4.4.1 below.

4.1.2 Technology Adoption in Developing-World Agriculture

Biochar appears to offer the prospect of improved agricultural productivity, both directly and via intensification through the improvement of returns to inorganic fertilizer. This prospect is particularly attractive given both that biochar may contribute to climate change mitigation, and that this mitigation potential might justify carbon finance to support dissemination. However, diffusion and adoption of new agricultural technologies has proven to be a slow and uncertain process, particularly in sub-Saharan Africa[2, 3]. While profitability of new technologies has been shown to often not be as great as is commonly supposed [6, 7, 25], demonstration of economic benefit where it exists is insufficient to guarantee rapid dissemination and widespread adoption[143].

Despite several decades of scholarship on the topic, it remains unclear how to stimulate dissemination efficiently. Early canonical papers tended to focus on the long-term patterns of technology diffusion, which they characterized using logistic models – slow early dissemination among innovators, followed by acceleration as the innovations caught on with more farmers, followed by a deceleration as more recalcitrant farmers begin to adopt[144, 145]. While these works remain touchstones in the literature, their key insights provide relatively little guidance to actors seeking to speed the rate of diffusion.

From independence to the 1980s, subsidies were the major tool used to stimulate diffusion of hybrid seed and inorganic fertilizer in sub-Saharan Africa. Faced with criticism for being inefficient, costly, and potentially regressive, these subsidies were largely phased-out during the structural adjustment era, which saw an end to the growth in fertilizer consumption seen in the region in the 1960s and 1970s[5]. During their heyday, fertilizer subsidies were justified using some of the main hypotheses driving empirical technology adoption literature in more recent years[146, 147], particularly that provision of subsidies would encourage learning-by-doing, and thereby stimulate demand at market prices[148]. While blanket subsidies tend to be poorly targeted and potentially inefficient[149], it is not always clear what alternative programs might work better. At minimum, the change in adoption with increasing subsidy can be used as a baseline with which to measure the efficacy of alternative measures.[5]

The fact that subsidies are costly has led to their justification on grounds other than simple dissemination; including the learning-by-doing argument mentioned above, as well as concerns for the welfare of poor farmers [150]. However, substantial research has also focused on whether demand might be stimulated in more cost-effective ways – potentially freeing resources for programs that might address welfare more directly. In particular, I focus on (social) learning and on novel sales offers designed to simultaneously remove risks and alleviate liquidity constraints. If equal rates of dissemination can be achieved

[5] Recent years have seen a resurgence in agricultural subsidies under the umbrella of "smart subsidies," which seek to improve targeting and efficiency as compared to experiences of post-independence period[150].

by stimulating social learning or providing these novel offers, as compared to subsidy programs, then such programs should be considered preferable as they would free resources for other development priorities.

Social learning and technology adoption has been a popular topic in the field for the last two decades. Beyond interest in furthering economic theory – social learning is notoriously difficult to distinguish from imitation and correlation[151] – interest is often motivated by the observation that technologies that achieve high rates of penetration weren't generally subsidized and seemed to spread virally, with the rapid spread of cellular phones in the region commonly serving as an example. Several recent studies, often based on field experiments, have measured the effect of adoption among social network contacts on own-adoption. Most find large effects [152–156] – on the order of 20% increase in adoption rates among those with treated contacts or neighbors, though others find more modest effects[143, 157, 158]. In particular, the relationship between adoption and the number of social network contacts with a technology has sometimes been found to be concave – social effects have been strongly positive when there are few adopters in a network, before declining and becoming negative at high levels of adoption – though different studies find turning points at different locations [153, 155, 157]. What's more, several of these studies have found that social effects work in some contexts but not others. Munshi found that social effects were stronger stimulating uptake of hybrid wheat than hybrid rice, arguing that wheat-producing areas were more environmentally-homogeneous than rice-producing areas, making learning easier in the presence of imperfect information about the mediators of the new technology's benefit. Krishnan and Patnam find substantially stronger social effects for seed adoption than for fertilizer adoption. Conley & Udry find that the social effects of "bad news" about a combination of inputs are stronger than those of "good news" about a combination of inputs. Finally, Duflo, Kremer & Robinson find that social effects are strong in the first year, but disappear in subsequent years – which is the opposite of what was found in Krishnan and Patnam's context, where the effects of extension[6] faded over time while social effects persisted.

Some of these studies distinguish between social learning via transmission of information about benefits of technologies, and social influence via imitation and herding[160], while others haven't. In general, a correlation between adoption in a network and one's own adoption could come either from learning or from imitation, in contexts where it is possible to exclude the possibility that results are driven by correlated unobservables. This paper focuses simply on propensity to adopt a technology as a function of the number of social network contacts that have adopted that technology. In a companion paper[7], I distinguish

[6]With exceptions[155, 159], studies of social learning and technology adoption in agriculture have dealt with traditional agricultural extension in little more than passing. This is largely a function of the developing-country context of the work, where extension services often have insufficient reach, and "model farmer" models seek to disseminate innovations to wider populations through first deploying the interventions to selected farmers, and then relying on social learning. Genius et al. find the two approaches to be complimentary, while Krishnan & Pattam find the extension has a much smaller effect than social learning. On the other hand, Maertens[157] finds that learning from authoritative sources is far more predictive of adoption than is the number of farmers in a network who have adopted.

[7]Companion paper in preparation.

between social learning and social influence, finding that social network effects operate primarily through learning.

For technologies with particularly heterogeneous benefits, learning from experience might be preferable to learning from others. Yet poor farmers may balk at the risk inherent in investing in a technology with unknown outcomes, and/or they may simply lack the liquidity needed to do so. Technologies may be offered in such a fashion that these potential barriers no longer bind. Given that such offers involve a cost to the entity sponsoring them, they can be considered preferable to a simple subsidy where the cost of provision is lower than the cost of the subsidy, for a given level of adoption within a population. In this study, I offer some farmers biochar for sale on a "risk-free trial" (R) basis – farmers receive a small quantity of biochar for a price, but are told that they can choose to pay for it or not at harvest. If they pay, then they can continue to purchase biochar in subsequent seasons. If they do not, then they will be excluded from future biochar sales. Similar work includes Levine et al.'s [123] work on improved cookstoves in Uganda. They find that the combination of time-payments (addressing liquidity constraints) with the right to return (addressing risk) the promoted stove led to 42 percentage points more adoption than control groups. While there isn't much literature on these sorts of offers, they represent the combinations of a number of mechanisms, each well-studied in development economics. While such offers may not make sense as a business model for sustainable provision of a technology in many contexts[8], they may represent a means of increasing demand by stimulating learning from experience, and it seems possible *a priori* that they could be more cost-effective in doing so than simple subsidies, if they are better at increasing uptake and are repaid at high rates.

4.2 Context, experimental design, and methods

I conducted this study in rural areas surrounding Bungoma, in Western Kenya. The region receives about 1600mm rainfall annually, distributed bimodally in a long and a short rainy season, with the former lasting approximately March to June and the latter from August to November. Agriculture in the region is dominated by maize and sugar cane production. The former is grown mostly for consumption, while the latter is a cash crop. Most farmers plant two maize crops per year, and yields in the short rains are typically half of those from the long rains. Soils in the region are relatively weathered[161]. Population density is higher than that in much of the rest of rural Africa, averaging 100-250 persons/km^2. Farm sizes average below 1 hectare, and fallowing – once common in the area – is now rarely practiced due to land scarcity.

The project started in early 2013 as a partnership with Re:char Kenya Ltd., a private business which sold simple biochar production equipment to farmers. Called the "Rutuba[9]

[8]Beyond the simple business risk in non-repayment, the threat of exclusion from future sales breaks down when more than one vendor is selling the technology in question.

[9]"Rutuba" means "soil fertility" in Kiswahili.

kiln," their device was a top-lit updraft gasifier[115] fashioned from oil drums and sheet steel. While the Rutubas are effective producers of biochar from agricultural waste such as maize stalks and sugar cane leaves, Re:char was unsuccessful in stimulating sufficient demand to sustain its operations, and went out of business in mid-2013. While a number of operational issues may have contributed to this outcome, a central factor explaining their exit may have been the focus on selling the technology itself – at a price of 4000 Kenyan Shillings (KSh) ($47.1) – which was felt to be too expensive by many in the region. Furthermore, former Re:char staff speculate that many of the kilns were adopted by farmers who had the intention of using them extensively, but who then did not do so during the periods when excess biomass became available, during harvest seasons. At Re:char's exit, my project acquired their stock of Rutuba kilns. Rather than selling the technology directly to farmers, I focused rather on selling the service of biochar production and application to farmers, as described below.

Before Re:char's entrance, some biochar dissemination work had been conducted by the African Christians Organization Network (ACON) – a small NGO based in Bungoma. ACON's work started in 2008, after good results were observed on small plots. ACON typically targeted self-organized community groups, who established biochar demonstration plots. After observing good results on these demonstration plots, training participants often moved on to use biochar on their own land.[10] Around the time of Re:char's entrance however, ACON's work shifted towards clean cookstoves – leaving my project as the sole purveyor of biochar in the region by mid-2013.

4.2.1 Experimental design

Because prior experience with either Re:char or ACON might influence biochar adoption, I selected areas around Bungoma where neither organization had previously worked. Specifically, I took a number of random draws from a circle centered on Bungoma, excluded those points that were within 4km of areas where either Re:char or ACON had previously sold kilns or done demonstrations, and selected the first 4 of those remaining. A 2km^2 box was drawn around these points, and households within the zone defined by the box were identified using satellite images. Figure C.1 gives the locations of the four zones, and the households within them.

A timeline of the project's activities is given in figure 4.1. Following approximate identification of households via satellite, enumerators visited these points, identified which of those represented households, and conducted the baseline census. Because I'm interested in agricultural decisionmaking, I define the "household" unit as groups of people who

[10] Most ACON-trained farmers used biochar from one of three sources: (1) collection of charcoal fines (small pieces and dust) from charcoal sellers in the area, (2) partially-combusted residues of home cooking fires, and (3) home production of cooking charcoal (commonly produced for sale), and utilization of the resultant dust and fine particles as biochar. None of these sources is readily scalable, and quantities used by these early adopters were quite small – a factor motivating Re:char's entrance.

farm separately. Families in this region are commonly polygamous, though some polygamous families farm separately, while others farm collectively. Thus a household could be a monogamous family, a polygamous family farming together, or one of several wives of a single husband who farm separately, together with the husband or separate from the husband, if he farms separately from his wives.

Beginning in August 2013, 75 farmers were selected at random to receive biochar demonstration plots.[11] Project staff asked participating farmers to identify a prominent 1/8th-acre portion of their farm – in the sense that it was relatively more visible to passers-by than other portions – to comprise the plot. Plots were divided into four portions – half was to get biochar,[12] while the other half didn't (which half got biochar was decided with a coin flip). Each of these two halves was further divided in half; we encouraged farmers to plant a staple food crop in one, and any alternative crop of their choice in the other. On the biochar half, project staff applied 25kg of biochar, which is equivalent to an application rate of 0.5 tons per hectare. Plots were demarcated using twine and marked with signs indicating which had biochar, and which didn't (figure C.2). Farmers were asked to treat both halves identically with respect to inputs and labor. The spatial distribution of the demonstration plots is given in figure C.3. At harvest in December, project staff measured yields on all demonstration plots using a scale.

We conducted a midline survey in October and November 2013, which focused heavily on social network data. For each respondent, I selected 75 other fellow respondent households to ask about, with the probability that farmer i was asked about farmer j being inversely proportional to the physical distance between them. For each potential link, farmer i was asked if they knew farmer j, using nicknames, spouse names, and place names to help distinguish people with similar names. Where a farmer i indicated that s/he knew j, surveyors asked about a number of different sorts of linkages, for example how recently they had seen one another, whether they talked about farming, etc. These are detailed in section 4.3.

Beginning in late January 2014, enumerators randomly allocated sales offers to all households. This was done by visiting all households in the the study areas, and offering a selection of sealed, unmarked envelopes, each containing an informational flyer and a sales offer with a randomized price. Under the "service model," farmers were offered the service of having our team visit their farm with Rutuba kilns, and produce char from their crop residues[13]. The service model was offered at randomly-varying prices, for fixed-ratio denominations of 50kg per 1/4 acre (i.e.: the application rate was controlled to remain roughly uniform across all who got biochar). In addition to the random prices offered to all farmers, some farmers were offered a "risk-free trial" (RFT), whereby they could get

[11]Not all of the original 75 selected to receive the plots; out of the original 75, 7 people refused, leading us to select replacements for them. Implications of this are explored in section 4.3, below.

[12]While biochar production can be somewhat labor-intensive and mediated by skill (depending on the technology used), biochar application involves very little skill or learning – it is applied together with fertilizers at planting, either in furrows or in small holes with seeds.

[13]Where sufficient crop residues weren't available, staff delivered biochar produced elsewhere.

a 1/4-acre biochar demonstration plot, but for which they would not be obligated to pay until after the subsequent harvest. If at that point they did not feel that the result was sufficient to justify the expense, they were not obligated to pay. If however they wanted to buy biochar from us again, they were expected to pay the price that was contained in the envelope. Payment was expected shortly after harvest – in order to minimize liquidity issues – but was accepted at any time. This distribution of prices and RFT offers is given in figure C.9. The maximum price – KSh1,250 (USD14.7) per 1/4 acre, or KSh625/sack (USD7.5) – was constructed with an approximate profit margin of 55% over the approximate production cost of KSh400/sack (USD4.7) to simulate the margin for a biochar vendor. During this randomization surveyors also asked about the number of biochar demonstration plots observed, and the respondent's subjective impression of the relative yield on the biochar plot as compared to the control plot.

The final survey round began in August 2014, in which we conducted a second sales offer. As I'll describe in section 4.4, sales in early 2014 were low. I therefore changed the sales model in the second round. First, farmers were offered biochar in 25kg sacks (sufficient for 1/8 acre), which were delivered and applied for a random price, ranging from KSh50-400 (USD 0.6-4.7), in increments of KSh50.[14] Unlike in the previous season, distribution of different subsidy levels was flat – more farmers were offered heavy subsidies. In addition, all farmers who had gotten a RFT in the previous season were offered a randomized discount on the cost of their RFT repayment, ranging from 50-90%.

Project staff weighted the harvests for all farmers who had gotten biochar for the 2014 long rains (in early 2014), from both the biochar portions, and the portions without biochar. In addition we collected plot-level data on fertilizer use, land size, and "selection" – asking farmers if the place where they applied the biochar was relatively less fertile, more fertile, or about the same as the portions of their land that did not receive biochar.

4.2.2 Semiparametric statistical methods

My analysis involves relating continuous treatment variables to non-Gaussian outcome variables. To avoid biases stemming both from functional form and distributional form mis-specification, I extensively employ generalized additive models. Generalized additive models (GAMs) [66] are semiparametric generalizations of generalized linear models (GLM's). The general form is $g(\mu) = \boldsymbol{W'\theta} + \mathcal{F}(\mathbf{Z}) + \epsilon$ where μ is the mean of y, which follows some distribution in the exponential family, $\boldsymbol{W}$ are variables associated with parametric slope coefficients and $\boldsymbol{Z}$ are variables represented non-parametrically. The function $\mathcal{F}(.)$ is commonly a sum of univariate smooth terms, but can also include smooth functions of more than one variable[162]. These smooth functions are typically represented by penalized splines; specifying $\mathcal{F}(.)$ terms using some spline basis – typically

[14]Rather than simply selling sacks of biochar, staff delivered them and applied them to soil as directed by the purchasing farmers. This was done to avoid the potential that a farmer who had randomly selected a low price could buy and resell to his/her neighbors at a higher price.

choosing a higher basis dimension than is needed (I use 10 per variable) – and choosing $\hat{\boldsymbol{\beta}}$ to minimize $-l(\boldsymbol{\beta}) + \sum_{i=1}^{m} \lambda_i \boldsymbol{\beta'S_i\beta}$, where $l(\boldsymbol{\beta})$ is the log likelihood of the model, $\boldsymbol{\beta}$ is here defined to include both $\boldsymbol{\theta}$ and the penalized coefficients associated with each spline term. The smoothing penalties λ_i control over-fitting – common to un-penalized spline models – and the matrices $\boldsymbol{S_i}$ are constructed so that smoothing only applies to non-parametrically represented terms. The λ_i are chosen to minimize bias while avoiding overfitting. I use the the restricted maximum likelihood smoothing parameter selection criterion[53] throughout, which is similar to the treatment of smoothing parameters as shrinkage factors in random effects models. Other approaches exist, including generalized cross-validation – these methods choose λ in order to minimize prediction error, but are sometimes prone to undersmoothing. When $\lambda \to \infty$, the result is equivalent to a generalized linear model (or OLS in the Gaussian case); GAMs include these as a special case. Where λ is constrained to be 0, the result is an un-penalized spline model, which will typically be overfit, particularly where basis dimension is sufficiently high as to avoid substantial bias. Where λ is estimated from the data however, the result is a flexible model in which both the functional form and the degree of smoothing are data-driven. When continuous variables have nonlinear effects, GAMs can be clear improvement over parametric techniques; reducing bias compared to linear specification, and often using fewer effective degrees of freedom than polynomial expansions. Where effects truly are linear, smoothing parameter selection algorithms will typically recover this. Textbooks on GAM's include [52, 163].

Where variables are endogenous – the case for previous season biochar adoption in the second season – I follow Marra & Radice's extension of instrumental variables techniques into the GAM framework, using two-stage generalized additive models (2SGAM)[164]. This framework itself is an extension of the "control functions" framework developed by Heckman[165], and is similar to the nonparametric instrumental variables methods developed by Whitney and Powell [166]. Briefly, the approach adjusts for endogeneity in variables $\boldsymbol{X}^{en}$ (defining $\boldsymbol{X} \equiv [\boldsymbol{W}, \boldsymbol{Z}]$) by controlling for smooth functions of the residuals of first stage regressions $g(x_p) = f(\boldsymbol{X}^{ex}, \boldsymbol{I}) + \nu$, where p indexes endogenous variables, and I are the instruments. In the second-stage regressions, I fit $g(y) = f(\boldsymbol{X}^{ex}) + f(\boldsymbol{X}^{en}) + \sum_p f_p(\zeta_p) + \epsilon$, where $\boldsymbol{X}$ includes exogenous and endogenous variables, and where ζ_p is a vector of residuals from the pth first-stage regression.

To account for uncertainty in the first stage regressions, the asymptotic variance-covariance matrix is calculated with a posterior simulation algorithm. Again following Marra & Radice, I first draw 25 simulates of the coefficients from the posterior of each first stage regression, treating the coefficients as $\boldsymbol{\beta} \sim \mathcal{N}\left(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\sigma}}_{\hat{\beta}}\right)$, yielding 25 draws from the distribution of ζ_p for each p. Secondly, the second-stage model is then re-fit with each of these 25 simulates, yielding 25 draws of the second stage coefficients [15]. Third, 100 simulates

[15] This procedure is implicitly Bayesian, though a frequentist interpretation might be that we are drawing different estimates of the $\boldsymbol{\beta}$'s that we'd get had we sampled differently from the wider population.

are taken from the distribution of the coefficients from each of these second-stage models, yielding 2500 draws in total. The asymptotic variance-covariance matrix is simply the covariance of these coefficient draws.

4.3 Data and descriptives

Summary statistics on the farmers in the sample are given in table 4.1. The sample is composed primarily of small farmers with low yields and little other income. Women are the main agricultural decisionmakers in about 2/3rds of households. Few say that they had ever heard of ACON, Re:char, or the Rutuba kiln, and few (5%) report ever using charcoal as fertilizer. Fertilizer use is generally low – most farmers spend just over $78/ha/season, while the government recommendations for maize in the region call for 125kg of di-ammonium phosphate (DAP) and 125kg of calcium ammonium nitrate (CAN) per hectare, which together tends to cost around $200.

4.3.1 Compliance and attrition

Of the 1115 people surveyed at baseline, 82 were selected to receive demonstration plots, 9 of whom refused[16]. Five of these refused because they did not plan to farm in the 2013 short rains. The maize on one demonstration plot was eaten by a cow, leaving 72 total demonstration plots. Summary statistics disaggregated by demonstration plot status are given in table C.1. Those who got the demonstration plot are somewhat different from those who did not, though these differences aren't generally statistically significant. There are however marginally significant differences in terms of bicycle ownership and having heard of the Rutuba – a factor that I control for throughout. The 9 farmers who refused were more likely to own a bicycle than other farmers in the sample (all had bicycles), and had more land under beans – a result driven by a single outlier.

Over the course of the project after baseline, 289 farmers attrited at any point during the project – they either refused to be surveyed or could not be found by the survey team during one of the subsequent rounds. Many of these were missed in one round but included in later rounds. For example, 873 farmers received sales offers in early 2014, and 900 received sales offers in late 2014. Statistics from the baseline census by attrition (defined as having ever missed either a sales offer or a survey round) are given in table C.2. Those who attrited tend to have similar characteristics than those who did not, except that those who attrited are less likely to have a male as the main agricultural decisionmaker, somewhat less likely to be listed as a social contact, and have somewhat higher incomes – generally consistent with households who have members who travel

[16] These refusals include 7 refusals from the original pool of 75 demonstration plot draws, and 2 refusals from the 7 replacements selected subsequently.

outside of the region semi-permanently for work. I control for each of these factors throughout.

4.3.2 Social networks

Social network statistics are given in table 4.2. The average farmer knew about one third of the 75 that were asked about, but talked about farming with far fewer. The average farmer had seen the land of about 1/5th of those asked about.

I don't assume that a farmer who says that they know another farmer is known by that farmer – I assume a directed network. While the literature is dominated by analysis of undirected networks, assuming an undirected network is problematic in this context – of the 7953 instances where both farmer i was asked about farmer j and vice versa, the link was reciprocal only 67.6% of the time.

Inferences based on networks constructed from sampled nodes are generally biased[167]. I avoid this bias by imputing missing edges using their conditional expectations, predicted from the following mixed logit:

$$pr(link_{ij}) = \Lambda \left[\alpha + \begin{bmatrix} \text{Same church}_{ij} \\ \text{Same school}_{ij} \\ \text{Gender}_i \\ \text{Gender}_j \\ \mathbf{1}\left(\text{Gender}_{i==j}\right) \end{bmatrix}' \boldsymbol{\gamma} + (\beta + \beta_i + \beta_j)\, log(\text{Distance}_{ij}) + \epsilon \right] \tag{4.1}$$

Model 4.1 represents the probability of linkage as a function of distance, at which certain individuals know and are known at different distances, and controls for gender of the household's main agricultural decisionmaker, observed membership in common churches, and observed attendance of their children in common schools. I fit model 4.1 for several dimensions of social network linkage, and the results are given in table C.3. To varying degrees, distance[17] is a good predictor of social network linkage across all dimensions. Men report linkage more frequently, and are reported as known, more frequently than women. Farmers tend to know farmers of the same gender more often, people in the same church are more likely to know each other, as are people with children in the same school. These models are reasonably accurate – distributions of absolute prediction error (observed minus estimated probability) are given in figure C.4, and are all less than 25% on average. I use these models to impute links with their conditional expectation where missing, which are aggregated by respondent to calculate estimated links, given in table 4.2.

[17] Distance is represented in units of latitude/longitude, which are approximately equivalent near the equator, where Bungoma is located. 0.01 degrees of latitude is approximately 1km.

4.4 Results

Results are presented chronologically. I begin with demonstration plot yields, followed by biochar uptake in early 2014, followed by yield results from the 2014 long rains, followed by biochar uptake ahead of the 2014 short rains.

4.4.1 Late 2013: Crop yields on demonstration plots

Yield results on demonstration plots are shown in figure 4.2. On average, biochar plots performed substantially better than non-biochar plots. Assuming a maize price of \$392/ton (the approximate average retail[18] price during the project), biochar is profitable in expectation at a cost up to \$125/ha (KSh1083/quarter acre) for the average yields seen in these demonstration plots, and this expectation is positive at 95% confidence at a cost up to \$98.7/ha (KSh850/quarter acre)[19]. This is less than the cost of the most-expensive randomly-allocated price of KSh1,250 (\$14.7) per 1/4 acre, which is equivalent to \$145.3/ha.

Substantial variability in observed response to biochar suggests that biochar adoption might not be favored by risk-averse farmers. I explore this by determining what sorts of risk attitude are theoretically consistent with biochar adoption being favored or disfavored at a price of KSh850 per quarter acre, which is the highest price for which the expectation of biochar's profitability is significant at 95% confidence. I model the utility of adoption using a mean-standard deviation formulation following Saha [168], where $U(\mu, \sigma) = \mu^\theta - \sigma^\gamma$, and where the two exponential parameters control the degree of preference for distributions of different shapes. From the demonstration plots, average profitability is \$27/ha, with a standard deviation of \$140. Utility across a range of θ and γ is given in figure C.5. At KSh850 per quarter acre, moderate levels of risk aversion may make biochar adoption an unfavorable prospect for those who might be characterized by declining average risk aversion[20] with respect to this particular decision.

Risk may be an important reason why a farmer might disfavor the prospect of biochar adoption, as it is quite profitable in expectation. It is unlikely that there are substantial hidden costs. Many novel agricultural technologies are labor-intensive, which involves cost beyond the cost of purchasing the technology[169]. Biochar does not require any additional labor for application beyond that which is used to apply fertilizer. And while

[18]The majority of the farmers in the sample consume the majority of the maize that they produce; very few sell in denominations larger than the 90kg sack. I therefore use retail prices for 90kg sacks, rather than bulk grain prices.

[19]This is calculated simply by finding the highest biochar price for which the t-statistic of the difference in in expected revenue between biochar plots and non-biochar plots – using the \$392/ton maize price assumption – is greater than 1.96.

[20]Saha shows that $\theta < 1$ corresponds to declining average risk aversion, and that $\theta > 1$ corresponds to increasing average risk aversion.

any yield-increasing technology will increase weeding, harvesting, and processing labor requirements, it is also true that biochar's benefits in terms of durable yield increases and improved responsiveness to fertilizer are imperfectly observable. As such, it is not possible to exhaustively characterize all unobserved farmer-specific characteristics that might make biochar unprofitable, though it does not seem that biochar suffers from substantial hidden costs.

I explore mediators of observed demonstration plot results (for maize only) by fitting models of the form

$$\frac{\text{Biochar yield}}{\text{Control yield}} = g^{-1}\left(\begin{bmatrix} \text{DAP fertilizer} \\ \text{CAN fertilizer} \\ \text{Urea fertilizer} \\ \text{Control yield} \end{bmatrix}' \boldsymbol{\beta} + \epsilon\right) \tag{4.2}$$

where $g(.)$ is identity or log, with Gaussian or gamma disturbances, respectively. Results are given in table 4.3. Unconditionally, the average biochar plot yielded 37% more than its control. On these demonstration plots, biochar increases response to fertilizer – each additional kg/plot of calcium-ammonium-nitrate (CAN) fertilizer predicts a 25 percentage point increase in the yield ratio, while each additional kg/plot of di-ammonium phosphate (DAP) predicts a 13 percentage point increase (interactive effects of the two are negative, likely reflecting the fact that both supply nitrogen). Controlling for fertilizer, biochar does better on lower-yielding farms – each ton of maize per hectare reduces the expected yield ratio by 32 percentage points. I also fit a log-linked gamma GAM representing yield ratio as a function of control yield and fertilizer expenditure, which I calculate from reported fertilizer application rates using market prices prevailing at the time. Marginal effects from that model are given in figure 4.3 – each 1000KSh increment in fertilizer expenditure predicts a 10 percentage point increase in yield ratio at low expenditure, slowing at higher expenditures.

When asked, most farmers who hosted demonstration plots reported that the yields on the biochar portion were "very good" or "good" compared to the non-biochar portion of their plot (figure C.6). When asked about the results of the demonstration plots on their neighbors farms, most farmers reported that they didn't know/see any demonstration plots. Those that did tended to say that the results were good or very good. However, there is little correlation between reported qualitative impressions and measured relative increases, either for demonstration plot holders assessing the performance of their own plots, or for others assessing the demonstration plots of people that they know (figure C.7), a finding that is largely invariant to which dimension of the social network is examined (table C.4).

4.4.2 Early 2014: Biochar uptake

Of the 868 farmers who took an envelope containing a price and a sales offer in early 2014,[21] 161 adopted biochar (18.5%). The adoption rate was 31.6% for those with a demonstration plot, and 63.9% for those with a RFT offer. Of the 72 with demonstration plots, 19 (26%) had RFTs. Only 16 farmers adopted out of the 638 that did not get a RFT offer (2.6%), of whom 4 had demonstration plots. Of those that did pay upfront, the average price was KSh293.75, which is a discount of 76.5% off of the maximum price of KSh1250 per 1/4 acre.

There is little bivariate correlation between biochar uptake and most social network measures. Figure C.8 gives bivariate logistic fits of uptake as a function of the proportion of each farmer's network with a demonstration plot, and of the number of links with a demonstration plot, controlling for the number of network links (following Kremer & Miguel [170]). Other dimensions of social network linkage are similarly non-predictive of uptake (table C.5). Two exceptions are savings groups and church groups; simple logits predict that having a larger proportion of one's church or savings group with demonstration plots is predictive of one's own uptake. I'm not able to exclude however that these results arise due to the multiple comparisons problem, as they are not robust to a Holm-Bonferroni correction.

I fit several specifications of the following model, representing adoption as a function of the treatments;

$$pr(\text{adopt}) = \Lambda \left[\alpha + f \begin{pmatrix} \text{Demo plot} \\ \text{RFT} \\ \text{Price} \\ \text{Share of network w/ Biochar} \end{pmatrix} + \boldsymbol{X'\beta} + \epsilon \right] \tag{4.3}$$

where $f()$ specifies its constituent terms with various interaction structures across specifications, and $\boldsymbol{X}$ is a vector of controls including having heard of Re:char and ACON, the gender of the household's main farmer, household income, whether the household reports having used charcoal as fertilizer before, zone dummies, and test a number of interactions between the continuous treatments. Full model output is given in table C.6 and figure C.10. As with the bivariate analyses, uptake is predicted by lower prices, being offered a RFT, and having a demonstration plot. The only significant interaction is between price and being offered biochar on a RFT basis. Estimates of the effects of continuous treatments (price, its interaction with the RFT offer, and the share of the network with a demonstration plot in the previous season) are associated with high smoothing parameters, and as such their effects are near-linear. Having had a demonstration plot is associated with increased propensity to adopt by a factor of 4, or a 0.10% increase in

[21] Actually, 873 farmers took envelopes with sales offers, but I drop 5 observations with incomplete or missing data in any of the covariates used to fit the model, below.

average probability[22]. Being offered a RFT rendered farmers approximately 439 times more likely to adopt, or a 57 percentage point increase in probability of adoption. For those who were not offered RFTs, increasing price by a \$1 (∼KSh87) increment approximately halves (0.47) propensity to adopt, or decreases average adoption probability by 12.5%. Those who were offered RFTs were still somewhat price-sensitive, but less so; each \$1 increment added to the price decreased adoption propensity by approximately one third (0.68), or 5%. The effect of the share of the social network with biochar was not significant, either alone or interacted with any of the other exogenous treatments.

These results are not driven by anchoring around neighbors' prices, or the share of neighbors who got a RFT – figure C.12 gives bivariate nonparametric logit fits of the probability of adoption as a function of the average price in a network, and the share of the network that got a RFT. Neither measure is a significant predictor of uptake.

For comparison, I provide an equivalent OLS regression, a parametric logistic regression, and a Gaussian additive model in figure C.11. While results are qualitatively similar, there are some differences. In this context, F- and χ^2-tests indicate significant and insignificant ($p < 0.001, p = 0.47$) difference between the parametric and nonparametric Gaussian and binomial (logistic) models, respectively. As is common with Gaussian approximations of binomial distributions, these models give predictions outside of the [0,1] interval. In this context, estimates are reasonably robust to linear functional form assumptions and Gaussian distributional form assumptions. This will not be the case with analysis of late 2014 adoption, given below.

4.4.3 Mid 2014: Biochar impact on maize yields

I have two sources of data on the impact of the biochar that was adopted in early 2014. First, staff directly measured yields (i.e.: with a scale) on both biochar-containing and no-biochar portions of fields of adopting farmers, along with the precise extent of their plots under maize, using GPS devices. They then asked them how much fertilizer they applied to each portion of their land, and whether the place to which they added biochar was "more fertile," "less fertile," or "about the same" as the parts without. Second, I have self-reported data for all farmers. The two datasets have different strengths and weaknesses. The former avoids the inaccuracy inherent in self-reported data, whether

[22]Average marginal effects, along with marginal effects at the mean of covariates, can be misleading when there is more than one treatment – the case here. Because logistic regressions are not additively separable, the impact of the considered treatment – the one for which marginal effects are desired – is dependent on the values of the other treatments. Where there is only one treatment, the sample is a random draw from the population, and only one factor is exogenously manipulated, then the mean of the non-manipulated factors should be comparable to the population mean. Where there is more than one treatment, this is not the case. As such, average marginal effects *in the sample* cannot be interpreted as expected marginal effects *in the population*, unless non-considered treatments are held to zero. This is the approach presented throughout this paper. Estimated odds ratios, however, are preferable inasmuch as they are valid across individuals with heterogeneous inherent propensities within a given population.

from poor short-term recall, error, or untruthfulness, though it is likely that there are variables which vary between plots for which I lack data. The latter may be affected by error in recall or deliberate untruthfulness, though its structure lends itself to clean identification of causal effects of biochar adoption via instrumental variables. I'll use both methods, beginning with the former.

4.4.3.1 Direct measurements

On average, biochar plots were 46% the size of non-biochar plots, got 25% less fertilizer, and were usually less fertile (49%) or of similar fertility (50%) as non-biochar plots. The approximate cash value of self-reported inputs on biochar vs non-biochar plots was $32.8 vs $78.6 per acre. Yields on biochar plots compared to non-biochar plots are given on the left side of figure 4.4. On average, biochar plots yielded 0.5 ($\pm$0.25) tons/hectare more than non-biochar plots, with the difference being stronger where the non-biochar plots yielded less. This difference is higher at low yields and becomes insignificant where yields pass 2 tons/ha. Assuming a maize price of $392/ton (the average during the project), expected profitability of a 0.5 t/ha increase in yield is about $153/ha, above the highest of the randomly-allocated prices that was offered.

Given that farmers with biochar didn't manage biochar plots identically to non-biochar plots, this is likely a biased estimate of biochar's effect. I adjust for differences in how the plots were treated by fitting models relating yields to observed covariates, and then using these models to estimate what yields would have been had farmers used the same amount of fertilizer on both plots, the devoted the same amount of land to biochar as to control, and used the same quality of land. I begin by modeling

$$log(yield)_{ip} = \alpha_i + \sum_j f_j(\boldsymbol{X}_{ipj}) + \epsilon_{ip} \tag{4.4}$$

where p indexes plots (biochar or non-biochar), where $\boldsymbol{X}$ is a set of j explanatory variables including biochar use (as a dummy) interacted with DAP, CAN, and urea fertilizers, plot size, a factor variable for reported relative plot fertility. I then calculate adjusted yields as the fitted values of this model for mean levels of each continuous covariate, plus the estimated residual, with the biochar dummy equaling 1/0 respectively. Adjusted data is plotted on the right side of figure 4.4. This adjusted data is non-constant with respect to inputs and land size – the estimated relationship between biochar and non-biochar yields would be different at different covariate values. However, at the means of the continuous covariates the adjusted difference between biochar and non-biochar plots is larger than it is in the raw data. The fact that adjusted data tends to imply larger differences than raw data suggests that the differences measured from the raw data are a lower bound to biochar's true effect.

As with the demonstration plots, most farmers report that they observed small or large positive differences between their plots with biochar and their plots without (figure C.13, left). However, these reported impressions do not correlate with either raw differences between plots (not shown) or adjusted differences (figure C.13, right).

4.4.3.2 Self-reported data

Estimating the causal effect of biochar adoption using self-reported data is more straightforward. I fit

$$\begin{aligned} yield &= \alpha + \beta BC + \boldsymbol{X'\theta} + \epsilon \\ yield_{it} &= \alpha_i + P_t + \beta BC_{it} + \boldsymbol{X'_{it}\theta} + \epsilon_{it} \end{aligned} \tag{4.5}$$

using cross-sectional and panel data (recall-based beginning in 2010), respectively. Because biochar adoption is endogenous, I instrument with the randomized inducements analyzed above: having a demonstration plot, the share of network contacts with a demonstration plot, price, RFT, and an interaction between price and RFT.

The estimates of biochar's effect are generally negative and not significant (table 4.4), in contrast to estimates from directly measured data. To explore this contrast, I plot measured yields against reported yields, as well as measured land size against reported land size, in the upper panel of figure C.14. There is little correlation between what farmers reported and what was measured. While some farmers plant a portion of their maize on plots removed from the home area (where the biochar plot tended to be), which was not measured, this does not drive results – many farmers report having substantially less land than that which was measured.

To further explore potential problems with self-reported data, I take advantage of an administrative error which saw several respondents asked the same questions multiple times. While several survey modules were duplicated, I focus on maize yields for the short rains of 2013 and the long rains of 2014. For each farmer who was given this module more than once, reported yield ranges (the difference between the maximum they reported and the minimum they reported) are given in the lower panel of figure C.14. While most differences are small – likely representing approximation error – average differences are 82% and 102% of average maize yields for in the 2013 short rains and 2014 long rains, respectively. The magnitudes of these different answers to identical questions are not correlated with treatment status (having a demonstration plot, having biochar in 2014, biochar price, etc), nor are they correlated with membership in NGOs such as One-Acre Fund. This suggests that this misreporting is either simple approximation error (albeit very large error) or strategic behavior that doesn't vary systematically with any facets of this project – perhaps different farmers had different ideas of what the enumerators "wanted to hear." Irrespective of potential explanations however, self-reported yield data is not trustworthy. And while much work in development economics relies on self-reported data, this result is not the first time that the accuracy of self-reported data has been called into question [171, 172].

4.4.4 Late 2014: Second round biochar adoption

The original experimental design emphasized adoption and impact in the 2014 long rains. However, uptake in that season was very low. It is possible that too-high prices masked effects of learning about biochar's benefits, either from own-experience in having had it, and from the experience of seeing it's impact on the plots of neighbors. These considerations guided the re-randomization of prices in August 2014, for the short rains agricultural season, as well as the offer of discounts for RFT repayment.

4.4.4.1 Risk-free trial repayment

Of the 161 respondents with RFTs, 19 repaid ahead of the 2014 short rains planting season. Propensity to repay is negatively correlated with repayment price and with previous-month income, but is uncorrelated with measured results on demonstration plots, and reported impressions of outcomes on demonstration plots (table C.7).

4.4.4.2 Short rains biochar uptake

Between August and October 2014, 96 farmers bought biochar. Unlike the previous season, nobody was offered a RFT. The average price at which biochar was purchased was KSh143 per 25kg sack, similar to the average price of KSh294 in the previous season, for 50kg. The two seasons are not directly comparable however; the short rains (in the second half of the year) are generally less important agriculturally. This suggests that demand should be lower than in the long rains, all else equal.

Various bivariate descriptive plots are presented in figure 4.5. A smaller proportion of adopters had demonstration plots or biochar in the previous season (figure 4.5a & c). As previous-season biochar is now endogenous, I instrument for it using fitted values of model 4.3; biochar adoption in early 2014 is associated with a lowered propensity to adopt in late 2014(figure 4.5d). The share with biochar in the previous season does not have a significant bivariate association with adoption propensity (figure 4.5e). It is endogenous, though instruments for it exist in the network analogs of predictors of individual adoption in the previous season – the share of the network with RFTs (figure 4.5f), and the share of the network receiving low-priced (KSh100-250 per 1/4-acre plot) offers (figure 4.5g)[23], and their interaction (figure 4.5h). However, the relationship between adoption and the instrumented share of the network (figure 4.5i) with biochar shows little difference from

[23] I also explored the average price offered in a social network as a predictor of the share of the network receiving biochar. There is a slight negative association between the share of the network adopting biochar and the average price in the network, though the effect is weaker than that of the share of the network receiving an cheap offer, as presented in figure 4.5g. Furthermore, given that adoption propensity was quite nonlinear in price (figure C.8), the share receiving offers for the lowest prices should theoretically be preferred.

the endogenous version – no significant association, but suggestive evidence of a concave relationship. Unsurprisingly, price (figure 4.5j) is a significant predictor of uptake, but the discounted repayment cost for those with RFTs is not associated with adoption.

I combine the above into a single model:

$$pr(adopt) = \Lambda \left[\alpha + \begin{bmatrix} \text{Demo} \\ \text{BC}_{t-1} \end{bmatrix}' \beta + f \begin{pmatrix} \text{Share with demo} \\ \text{Share got BC 2014} \\ \text{Price} \\ \text{RFT repay price} \end{pmatrix} + \boldsymbol{X'\beta} + \sum_i^2 g_i(\nu_i) + \epsilon \right] \tag{4.6}$$

where terms listed in red are endogenous, and functions $g(\hat{\nu})$ are functions of the residuals of the first-stage regressions of the endogenous variables on the instruments, added to adjust for endogeneity. Results are given in table 4.5. First stage models are highly significant. The price/RFT interaction predicts prior-year adoption, and the interaction between the network share having a cheap offer (KSh100 or KSh250) and the network share having a RFT is a highly-accurate predictor of the share of the network with biochar.

In the second stage, we see – unsurprisingly – a negative relationship between biochar price and propensity to adopt, with adoption propensity increasing – and slightly accelerating – as prices decline. A \$1 (~87KSh) reduction of the full price of \$4.6 (400KSh) increases the odds of adoption by a factor of 2.1, while a \$1 increase from the minimum price that was randomly allocated (\$0.58) decreases the odds of adoption by a factor of 4.4. In terms of average marginal effects (holding all other treatments to zero), these represent a 1.5% increase and 6.2% decrease in probability of adoption, respectively.

The estimated effect of having had biochar previously – either in a demonstration plot or in the previous season – is negative, though not statistically significant. (The point estimate suggests a decrease in odds of adoption by a factor of 2.4, with an average partial effect of less than 1% when other treatments are held to zero). The estimated effect of the share (or proportion) of the network having biochar is also not significant, though marginal at $p \simeq 0.116$. Increasing the share of the network from zero to the estimated optimum – 12% – increases adoption propensity by a factor of 13.7, though only 2% in absolute probabilities (again, assuming no subsidy or previous exposure to biochar). Intervals on the effect of the network share indicate high confidence that the effect of the social network does not increase beyond this optimum, and that it may decrease. This would be similar to findings by others in literature on social influence and adoption [153, 155, 157] who report concave relationships between uptake and the share of a social network with a technology. However, this apparent decline occurs over a region of the function with relatively little support.

For comparison, I fit model 4.6 as a two-stage least squares (2SLS) linear probability model (LPM), including quadratic terms where suggested by the nonparametric analysis. Estimates are given in table C.8. The LPM gives larger and more significant estimates than does the GAM – increasing the network share from 5% to 15% suggests a 10% increase in adoption propensity, and the two quadratic terms are jointly "significant" at $p \simeq 0.018$. However, the LPM my offer spurious optimism under the present circumstances. First, it will be more sensitive to outliers, and there is relatively little support at the parts of the distribution where strong positive and negative effects on adoption propensity are modeled by the quadratic. Second, the sample was heavily subsidized on average – these estimates can be viewed as generalizing to a meta-population that was similarly subsidized, rather than the population from which the study's sample was taken. In other words, the additive separability of the LPM renders it misleading.

4.4.5 Prospective impact of subsidies on welfare and carbon sequestration

The foregoing results do not directly compare the efficacy of funds spend subsidizing biochar versus funds spent stimulating social learning, either in terms of uptake, welfare, or social cost/benefit. Furthermore, they don't compare the efficacy of money spent stimulating biochar uptake (potentially via a combination of subsidies and demonstrations) and money spent subsidizing inorganic fertilizer. I do so in this section, by taking conditional expectations of model 4.6 (negative binomial) across a grid of values constructed to represent hypothetical biochar subsidies and hypothetical expenditure on demonstration plots, using costs specific to this project's context. I do so for a median farmer – using response ratios at median baseline yields, median fertilizer application rates, and median covariate values of model 4.6, etc. What follows is presented in the spirit of a "back of the envelope" calculation – conclusions will depend on simple modeling assumptions, though the implications of altering these modeling assumptions should be transparent.

First, I calculate the expected number of links that a randomly-chosen farmer would have if I had given out a different number of demonstrations. I do this by taking the estimated probabilistic adjacency matrix $\boldsymbol{A}$ representing the social network, and for a gradient of demonstration plot numbers between N=1 and N=200, take 500 samples (without replacement) of farmers $\boldsymbol{g}$, where each row is a vector of N 1's and the rest zeros, for each farmer in my sample. This gives 500 different possible combinations of how the demonstration plots could have been allocated for each N. I then calculate the expected number of links that a given farmer would have to other farmers with a biochar trial as the mean of the row-wise sum of $\boldsymbol{g'A}$.

I use this to calculate the conditional expectation of the number of bags of biochar purchased across a grid of subsidies ranging from free provision to full cost (KSh400/USD4.6 per 25kg bag), and number of demonstration plots ranging from 0 to 200. This is given in figure 4.6a. The function is maximized with approximately 12% of the population getting demonstration plots[24], where biochar is heavily subsidized or nearly free.

[24]The median farmer in my sample farms one half acre of maize. As such, I set the maximum biochar purchase at 4 bags, as each is sufficient for 1/8 acre.

Figure 4.6b gives program cost in USD/farmer. For provision of demonstration plots, I assume that each costs double the cost of the biochar itself – a rough estimate based on this project's implementation costs – and divide the sum of these costs by the number of farmers in the sample. Demonstration plot costs are absolute, while subsidy costs are conditional on uptake by farmers at that level of subsidy and number of demonstration plots. The blue dashed line in figure 4.6b gives the "optimal ridge" – the combinations of subsidies and demonstration plots that maximize expected uptake for a given program expenditure. The best estimate of optimal program expenditure given my assumptions would involve heavy subsidies at low per-farmer program expenditure, then lowering subsidies and increasing expenditure on demonstration plots as per-farmer program expenditure increases. Figure C.16 gives confidence regions for the "optimal ridge" estimate based on uncertainty in model 4.6. I cannot exclude (at 95% confidence) that the optimal program for the median farmer would spend zero on demonstration plots, though it is increasingly likely that demonstration plots make sense as program expenditures increase. By contrast, 95% of all scenarios in which per-farmer expenditure is above \$6 include subsidies above 50%.

Figure 4.6c gives the implications of the subsidy scheme for the median farmer's returns to maize farming net of inputs. This is calculated as

$$\frac{\left((1-p)(\bar{y}) + p(\bar{y} * \widehat{RR} - c^{bag} * bags)\right) - \bar{y}}{\bar{y}} * 100\%$$

where p is the proportion of the median farm's maize allocation (0.2 ha) covered by the number of purchased bags of biochar, $\bar{y}$ is median market-valued maize production (222kg at \$0.39/kg) less median market-valued fertilizer usage (\$15.5). Response ratio (RR) – the expected ratio of biochar yield to non-biochar yield for a given level of fertilizer – is estimated from the model given in figure 4.3 for median non-biochar yields and median fertilizer expenditure. Profit is maximized where adoption is maximized – at free provision with about 12% of the network having a demonstration.[25]

Figure 4.6d gives social return on investment, defined as the difference between farmer profits under the inducement scheme and the baseline, less the cost of those inducements (which are dependent on uptake). I find that social return on investment is maximized at subsidies that induce maximal uptake, topping out at 25%. Incorporating carbon finance changes the locations of the optima little, though it increases social ROI by as much as 22 percentage points at optimal uptake. To calculate this, I assume that biochar is half carbon by mass, that 80% of this carbon is "stable," and that carbon is valued at \$100/ton. These are generous assumptions – particularly on carbon price and on biochar's stability. However, even holding this stability assumption, biochar would

[25] Note that estimates at free provision involve extrapolation between the maximum subsidy that was experimentally provided – 87.5% – and 100%. While nonlinearities in response in this interval can't be excluded, the optimum with the rightward part of the figure removed still involves maximal subsidies and about 12% of the network having demonstrations.

require a carbon price of at least \$120/ton for its inducements to be justified solely on the value of sequestered carbon (figure 4.6f). In general, it appears that the impact of a carbon price on the attractiveness of expenditure on biochar subsidies is marginal given extant carbon markets[26], though biochar inducements could be justified purely on a welfare basis.

These estimates apply to a median farmer in the present dataset. In general, these estimates suggest that biochar will be increasingly favorable where food prices are higher, soils are more degraded, fertilizer use is higher, and farmers are more land-constrained.

4.4.5.1 Subsidizing biochar versus fertilizer

Holding fertilizer expenditure constant, the cost of biochar inducements should yield substantial social return. A natural question is whether better returns might be had elsewhere, for example in subsidizing inorganic fertilizer. While I didn't randomize fertilizer subsidies, nor do I observe sufficient market variation in fertilizer prices to estimate a demand curve, I can generate first-order estimates of biochar's relative benefit under the assumptions – perhaps unrealistic – that farmers would hold expenditure on inorganic fertilizer constant in the presence of less-expensive fertilizer, or that farmers would hold total input expenditure (biochar plus fertilizer) constant in the presence of subsidies for both inputs. I represent expected yields as

$$yield = p\left(\hat{f}_Y(\text{Fert. Exp.}) \times \hat{f}_{RR}(\text{Fert. Exp.})\right) + (1-p)(\hat{f}_Y(\text{Fert. Exp.}))$$

where $\hat{f}_{RR}$ is the ratio of biochar yield to non-biochar yield estimated from the model given in figure 4.3, and evaluated at the median yield in the long rains 2013 (before the commencement of project activities). Yields as a function of fertilizer are estimated from the fixed-effects model $yield_{it} = \alpha_i + P_t + f(\text{Fert. Exp.}) \times SR + \epsilon_{it}$ where SR is a factor variable equaling one during the short rainy season, and t indexes agricultural seasons[27]. Predicted yield response to fertilizer in the long rainy season is given in figure C.15.

Figure 4.7 gives estimated social return on investment over a gradient of biochar inducement expenditure and fertilizer subsidies, for the median farmer. The dashed blue line gives the point of optimal social return-on-investment for a given combined biochar and fertilizer program expenditure. Not accounting for the value of sequestered carbon, and under the assumption that fertilizer expenditure remains constant as prices increase (left), or that total expenditure on biochar plus fertilizer is held constant (right), fertilizer subsidies provide higher social ROI than optimal biochar inducements. While biochar and

[26] Though this estimate of \$120/ton is consistent with some valuations of the social cost of carbon[108], and lower than others.

[27] This model is fit with self-reported data, for which I suspect substantial misreporting in later periods. However, given that this model is fit entirely with data collected before commencement of project activities, I don't expect intentional misreporting to be a problem.

fertilizer are highly complimentary, the marginal effect of subsidizing fertilizer is substantially stronger than the interactive effect of biochar and fertilizer together, until fertilizer use increases substantially. Confidence regions for these estimates – based on posterior simulation of models predicting adoption, yield, and biochar response ratio – are given in figures C.17 and C.18. With increasing per-farmer expenditure, fertilizer subsidies provide positive social return with very high confidence. Optimal expenditure on biochar subsidies is less certain, and generally not different from zero at 95% confidence, though the probability that they are are worthwhile increases with total program expenditure.

It should be emphasized that these are first-order estimates, and suffer from a few shortcomings. First, the assumption that farmers would hold expenditure constant as prices decrease is dubious at best – it seems plausible that farmers could either increase or decrease their total expenditure on fertilizer and biochar given a drop in prices. If farmers increase fertilizer expenditures as prices decline, fertilizer would be further favored over biochar, unless biochar's interactive effect with fertilizer is higher than assumed here. Second, these estimates rely on an estimated relationship between maize yields and fertilizer that is based on questionable self-reported data, and which may suffer from omitted variables bias. If the true relationship is more concave (linear) than assumed here, biochar inducements would be favored sooner (later). Third, these estimates represent a snapshot in time. It remains to be seen how farmer uptake of biochar – along with farmer response to subsidies and in-network demonstrations – would change given more experience with the project, or in years with different characteristics. Fourth, the effects of biochar on yields – both directly and through its mediation of fertilizer's efficacy –are likely to be persistent in soil over time. As such, biochar's benefit is likely to be substantially underestimated here. Fifth, these estimates can only be interpreted in a partial-equilibrium sense. Any large-scale program subsidizing biochar and fertilizer is likely to increase production, lower local prices (particularly given poor integration with world grain markets), and thus lead to lower returns. Magnitudes of such potential effects are unclear in this context, and will vary substantially between contexts as well. Finally, these estimates are calculated for the median farmer in my Western Kenyan sample; results will vary substantially across contexts. In especially degraded soils, biochar may be substantially more useful than fertilizer, making the curves presented in figure 4.7 rise more slowly and move right more quickly. Conversely, among higher-yielding farmers with fertilizer-responsive soils, biochar may not be useful at all. Likewise, biochar may make more sense in countries with lower fertilizer prices, and less sense in contexts where biochar is more costly to produce.

4.5 Discussion

Altogether, I find that biochar has substantial agronomic utility, but that significant investment might be required to lead to large scale adoption – at least initially. After two seasons, demand is substantially lower in this context that what would be required to sustain a biochar-producing firm. Yet public investments in increasing this demand seem likely to lead to very high social returns.

This is based largely on biochar's impact on crop yields. Here, though, the evidence is somewhat contradictory. Where directly measured, biochar led to substantial crop yield increases. However, I find no effect (and negative point estimates) when analyzing self-reported survey data on yields, or self-reported subjective impression of yield. It is difficult to avoid the conclusion that there may have been substantial survey nullification. In other words, farmers may have strategically told enumerators whatever they felt might be advantageous, based on what they supposed that the priorities of the project might be. Anecdotally, enumerators reported that many respondents conflated the project with One Acre Fund (OAF) – a prominent NGO operating in the region that offers agricultural inputs on credit. Many farmers in the region believe that OAF repossesses assets in the event of non-payment, though this practice isn't consistent with OAF's policies. On the other hand, OAF doesn't generally collect on loans when correlated shocks – such as poor rainfall – occur. This fear of debt, conflation of us with OAF, and perception that we might not expect repayment if results were poor, could have contributed to the divergence between what was measured and what was reported. While perhaps plausible, this is little more than speculation.

Uptake results are equally contradictory, though in different ways. Of the interventions implemented, the risk-free trial offers were the most effective in stimulating uptake. They were generally not repaid however, even when farmers' measured outcomes to adoption were favorable. Limited focus group discussions with free-trial recipients suggest that farmers wish to repay the RFT and buy more, but are not prioritizing this ahead of other costs. It seems likely that some form of psychological factor drives this finding, perhaps involving present bias; repayment of a trial of biochar given in a previous season in order to purchase more biochar now may be experienced by a farmer simply as an increase in the cost of biochar now. Given that farmers are highly price-sensitive, low levels of repayment are perfectly natural when seen in this light. In other contexts, where respondents faced the choice of paying for or returning an item – a cookstove, for example, which can be physically repossessed [123] – farmers are paying to keep something that they value, rather than paying to maintain the option to purchase more of something that they value, which will generate returns at a later point in time.

I find no evidence that demonstration plots among social contacts led to more uptake in the first season, though there is evidence that having a demonstration plot made uptake more likely. This result reverses itself in the second season, where I find suggestive evidence of an inverted u-shaped function relating uptake to the proportion of the social network with biochar – similar to findings of others in the technology adoption literature. It is unclear why the two seasons were different in this respect, though it is possible that farmers are learning about one another's *decisions* with regard to adoption, rather than purely viewing their neighbor's *outcomes* to adoption.

Unsurprisingly, price was a significant predictor of uptake in both seasons, but farmers became somewhat less price sensitive in the second season than they were in the first. More farmers bought biochar in the second season (96 vs 16), and the estimated functions mapping price to probability of adoption changed from one that where farmers were unlikely to buy at any price, to one where the probability of adoption declined steeply with price. It is unclear what caused this change, and longer longitudinal studies are needed to elucidate these dynamics. It is possible that farmers's beliefs about likely outcomes to adoption become decreasingly diffuse over time, and it is possible that the existence of the RFT offers made other farmers less willing to purchase for cash upfront, though it can't be excluded that the difference is simply driven by season-specific unobserved factors.

Further work is needed to place these results into a more dynamical context – to what degree can a "big push" of subsidies over a few years sustain long-term adoption, and will learning reduce the need for subsidies over time? Can these results be generalized to different technologies in different contexts? How does price sensitivity change further with time? Given that most of the current state of knowledge around technology adoption is based on cross-sectional data, further longitudinal work represents an important opportunity both to improve our understanding and to guide agricultural development policy.

4.5.1 Welfare implications of biochar subsidization

In increasing yields substantially, sequestering carbon, but yet being underadopted, biochar is a prime candidate for programs aimed at stimulating its adoption. In this project's context, it appears that heavy subsidies maximize social benefit/cost per dollar, along with a small number of demonstrations. In addition, these estimates of social ROI are probably underestimates, given the longevity of biochar's agronomic benefits observed in other studies, as well as the mitigation value of sequestered carbon.

In the short run, subsidizing biochar alone may not be as effective as subsidizing fertilizer, though this will vary with soil characteristics, fertilizer prices, and the economics of biochar production in different places. Provision of biochar can markedly increase fertilizer efficacy, but not sufficiently in this context to clearly favor biochar subsidies over fertilizer subsidies. This result does not account for the persistence of biochar's agronomic benefit, for which I have no data here, and will be an underestimate where soils are particularly degraded and response to fertilizer is low. In such contexts, biochar could serve a role in accelerating agricultural intensification in the region, particularly if its ability to improve the efficacy of fertilizer is maintained over multiple seasons. In a region where fertilizer is commonly perceived to be risky and expensive, biochar might make fertilizer effectively cheaper and therefore less risky by reducing the amount required for a given yield increment.

Biochar is likely to be more expensive that other means of increasing soil organic carbon content (such as composting, or integration of trees into agricultural landscapes). However, it increases yields immediately, and decomposes much more slowly than other organic amendments. These are important advantages – the former from the perspective of generating interest among poor, present-biased farmers, and the latter from the perspective of donors, who may be attracted by the prospect that one "treatment" of biochar may have benefits over multiple seasons. If long-term yield benefits to biochar in smallholder agriculture mirror those that have been found in the agronomic literature[32], this property of biochar endows it with a natural "exit strategy" – a donor could support the cost of its provision once, generating returns over multiple seasons, before moving on to a different area. Furthermore, the durability of biochar's benefit may mitigate concerns that its use may not be appropriate everywhere given biomass constraints. In general, opportunity costs for biomass (to serve as biochar feedstock) are not zero, and it could be difficult in many areas to find sufficient biomass every year for biochar production. Much like the way in which agricultural lime is applied semi-annually, semi-annual use of biochar may allow for large areas of land to progressively become amended with it where it is infeasible large quantities of it to be produced at once.

4.6 Conclusion

I've conducted the first randomized controlled trial on the adoption and impact of biochar in smallholder agriculture, and found substantial agronomic utility, severely sub-optimal rates of adoption, and some evidence to guide technology diffusion policy. More research is needed to explain why adoption was so insensitive to such large observed results. It appears that relying on social learning to shortcut the high cost of subsidies is likely to be misguided, but that it might be able to play a complementary role in making subsidies more effective. Similarly, biochar on its own can increase crop yields, but should probably be viewed as a means of improving soil quality and returns to inorganic fertilizer, thereby stimulating sustainable agricultural intensification. Together, these results demonstrate substantial potential for biochar in African agriculture, they suggest means of promoting it, and as such could guide approaches for further applied research and program development in the area.

4.7 Figures and Tables

FIGURE 4.1: A timeline showing project activities and the region's agricultural calendar.

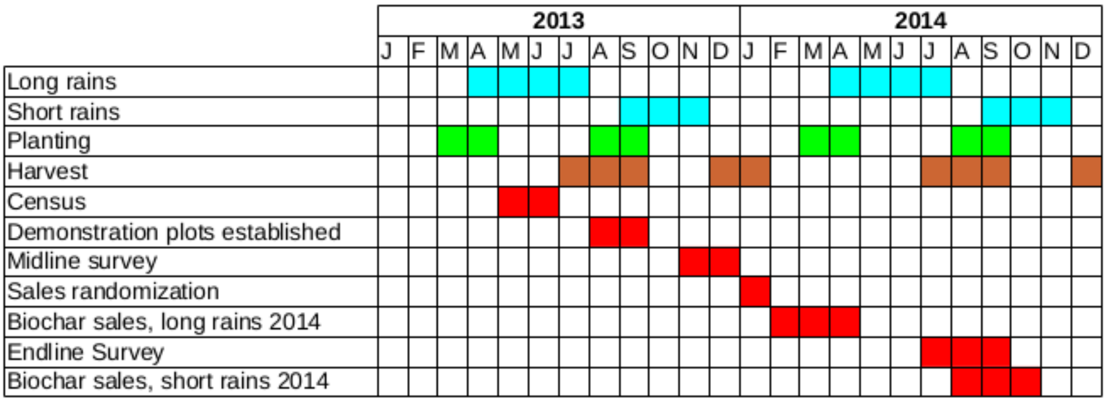


TABLE 4.1: Summary statistics describing the sample. Time-varying agricultural variables (yields, land allocation, etc.) taken from the long rains 2013, before any of the treatments were implemented.

Statistic	N	Mean	St. Dev.	Min	Max
Maize Yields (tons/ha)	864	1.093	0.81	0	3.34
Hectares Maize	864	0.35	0.320	0.05	3.24
Hectares Sugar	889	0.1	0.262	0	3.64
Total Land (ha)	889	0.56	1.752	0	51
Fert Exp ($/ha)	887	78.39	72.063	0	374
Income (Sep 2014)($/month)	895	61.79	108.32	0	1,458
Family size	911	5.89	2.49	1	16
in OAF	914	0.23	-	0	1
Gender main farmer (0=female)	914	0.35	-	0	1
Heard of Re:char	914	0.04	-	0	1
Heard of ACON	914	0.09	-	0	1
Heard of the Rutuba Kiln	914	0.051	-	0	1
Ever used charcoal as fertilizer	914	0.05	-	0	1
Ever produced charcoal	914	0.33	-	0	1

Table 4.2: Summary social network statistics.

Statistic	N	Mean	St. Dev.	Min	Max
Reported links	945	26.357	10.838	0	61
Estimated links	945	94.151	46.204	8.082	284.035
Rep. talked farming	945	7.180	8.375	0	55
Est. talked farming	945	20.613	26.468	0.258	237.993
Rep. seen land	945	14.520	9.818	0	59
Est. seen land	945	43.419	34.898	1.277	263.011
Rep. same church	945	5.608	5.849	0	29
Est. same church	945	20.765	20.445	2.124	119.939
Rep. same sav group	945	2.841	4.839	0	38
Est. same sav group	945	9.383	14.980	0.538	144.827
N within 1km	945	160.505	50.897	23	299

FIGURE 4.2: Crop yields on demonstration plots, compared to control plots.

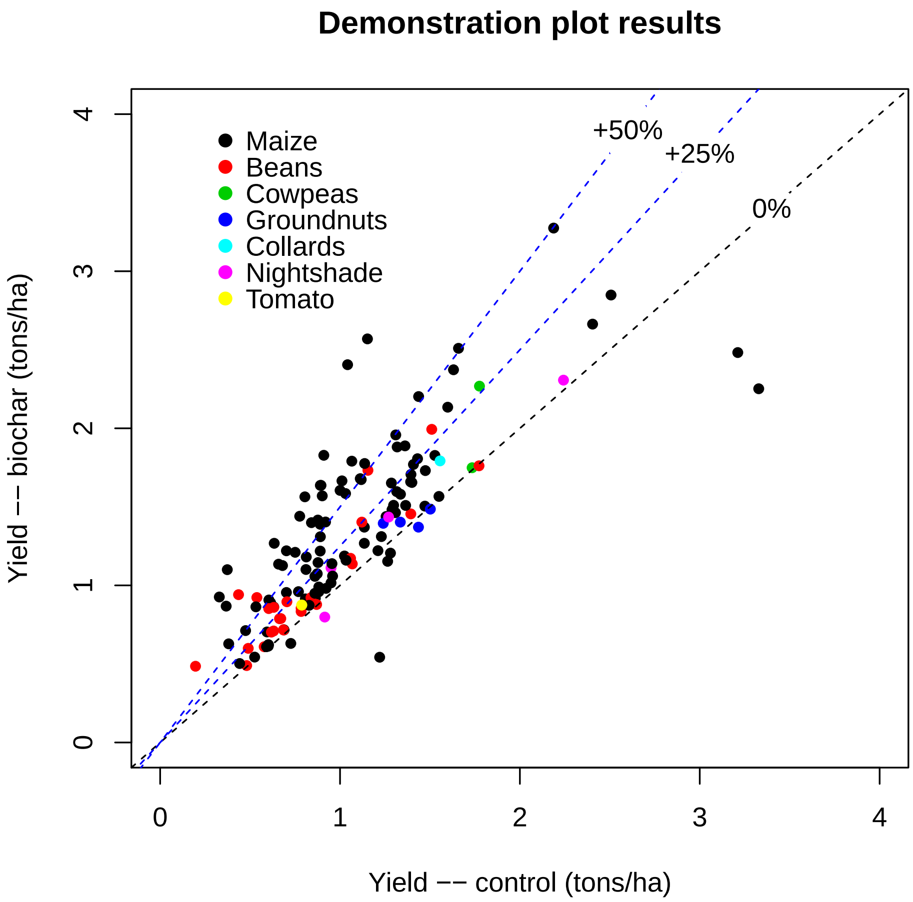


TABLE 4.3: Results of model 4.2 relating demonstration plot outcomes (biochar yield divided by control yield) to fertilizer use. Gamma results are raw coefficients – marginal effects on the link function (log) scale. Fertilizers are in units of kg/plot, while plots were 1/16th acre in extent. As such, each 1-kg/plot change represents a change of approximately 16kg/acre or 39.5kg/ha.

	OLS				Gamma			
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
(Intercept)	1.37^{***}	1.70^{***}	1.58^{***}	1.56^{***}	0.32^{***}	0.56^{***}	0.47^{***}	0.46^{***}
	(0.04)	(0.09)	(0.10)	(0.10)	(0.03)	(0.06)	(0.07)	(0.07)
Control yield		-0.26^{***}	-0.30^{***}	-0.32^{***}		-0.21^{***}	-0.23^{***}	-0.25^{***}
		(0.03)	(0.03)	(0.03)		(0.02)	(0.02)	(0.02)
DAP			0.07	0.13^{**}			0.05^{*}	0.10^{**}
			(0.04)	(0.04)			(0.02)	(0.03)
CAN			0.14^{**}	0.25^{***}			0.10^{**}	0.18^{***}
			(0.05)	(0.06)			(0.03)	(0.04)
Urea			0.04	0.03			0.03	0.02
			(0.05)	(0.05)			(0.03)	(0.03)
DAP×CAN				-0.12^{**}				-0.10^{**}
				(0.05)				(0.03)
AIC	102.31	88.21	80.52	75.28	87.85	69.67	61.77	54.62
BIC	107.52	96.02	96.09	93.45	93.06	77.49	77.34	72.78
Num. obs.	100	100	99	99	100	100	99	99

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Figure 4.3: Marginal effects of control yield and fertilizer expenditure (on a per-acre basis) on the mean ratio between biochar-plot yields and control-plot yields from the late-2013 demonstration plots. Yield ratio specified as gamma-distributed. Fertilizer expenditures calculated as costs of applied fertilizer at prevailing market prices.

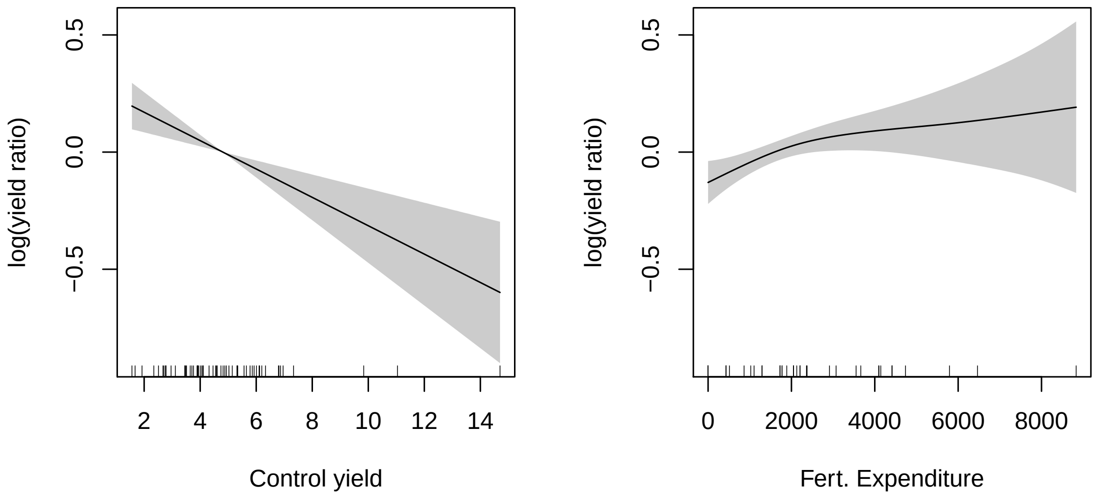


FIGURE 4.4: Maize yields on biochar plots as compared to non-biochar plots on fields of farmers adopting in early 2014. Left: raw data. Right: adjusted data, based on estimates from model 4.4. Univariate fits relating biochar yields to control yields are log-linked gamma GAMs.

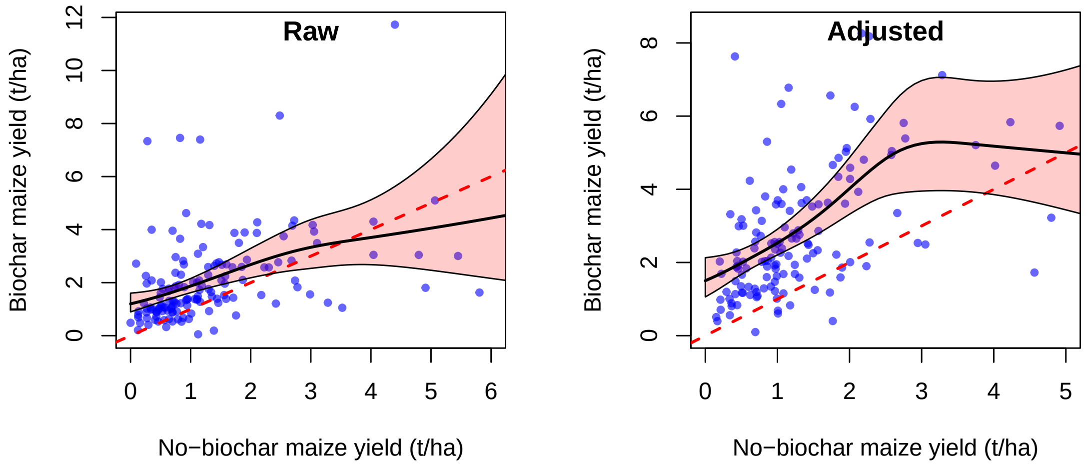


Table 4.4: Regression results from model 4.5; estimates of biochar adoption impact based on self-reported data.

	Dependent variable:							
	Maize yields (90kg bags/acre)							
	Cross-sectional					*Panel*		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Adopted Biochar	−0.304	−0.938**	−0.179	−0.717	−0.521	−0.339	−0.339	−0.176
	(0.326)	(0.373)	(0.462)	(0.562)	(0.510)	(0.313)	(0.313)	(0.287)
DAP fertilizer					0.028***			0.029***
					(0.007)			(0.005)
CAN fertilizer					0.047***			0.032***
					(0.008)			(0.004)
Urea fertilizer					0.024***			0.020***
					(0.008)			(0.006)
Acres under maize					−3.287***			−2.563***
					(0.311)			(0.208)
Constant	6.568***	60.288	6.545***	288.421	225.257			
	(0.135)	(493.732)	(0.155)	(488.757)	(441.718)			
Covariates	No	Yes	No	Yes	Yes	No	No	No
IV	No	No	Yes	Yes	Yes	No	Yes	Yes
Observations	865	860	792	792	792	5,821	5,821	5,808
R^2	0.001	0.132	0.001	0.099	0.268	0.567	0.567	0.619
Adjusted R^2	−0.0002	0.088	−0.0004	0.066	0.237	0.482	0.482	0.543

Note: *p<0.1; **p<0.05; ***p<0.01

FIGURE 4.5: Descriptive bivariate plots relating biochar adoption in late 2014 to prior-season and current-season treatments. Plot (d) sets probability of biochar adoption in late 2014 against estimated probability of adopting biochar in early 2014, estimated from model 4.3. Plots (f,g, and h are first stage regressions, predicting the share of the social network having biochar in the previous season as a function of the proportion with a RFT offer and the proportion with low biochar prices. Plot (h) presents a nonparametric interactive model, with the expectation of the share of the network on the z-axis. Plot (i) gives adoption as a function of the estimates of the model shown in plot (h). Plot (k) gives adoption propensity as a function of RFT repayment price, for those who had recieved RFT.

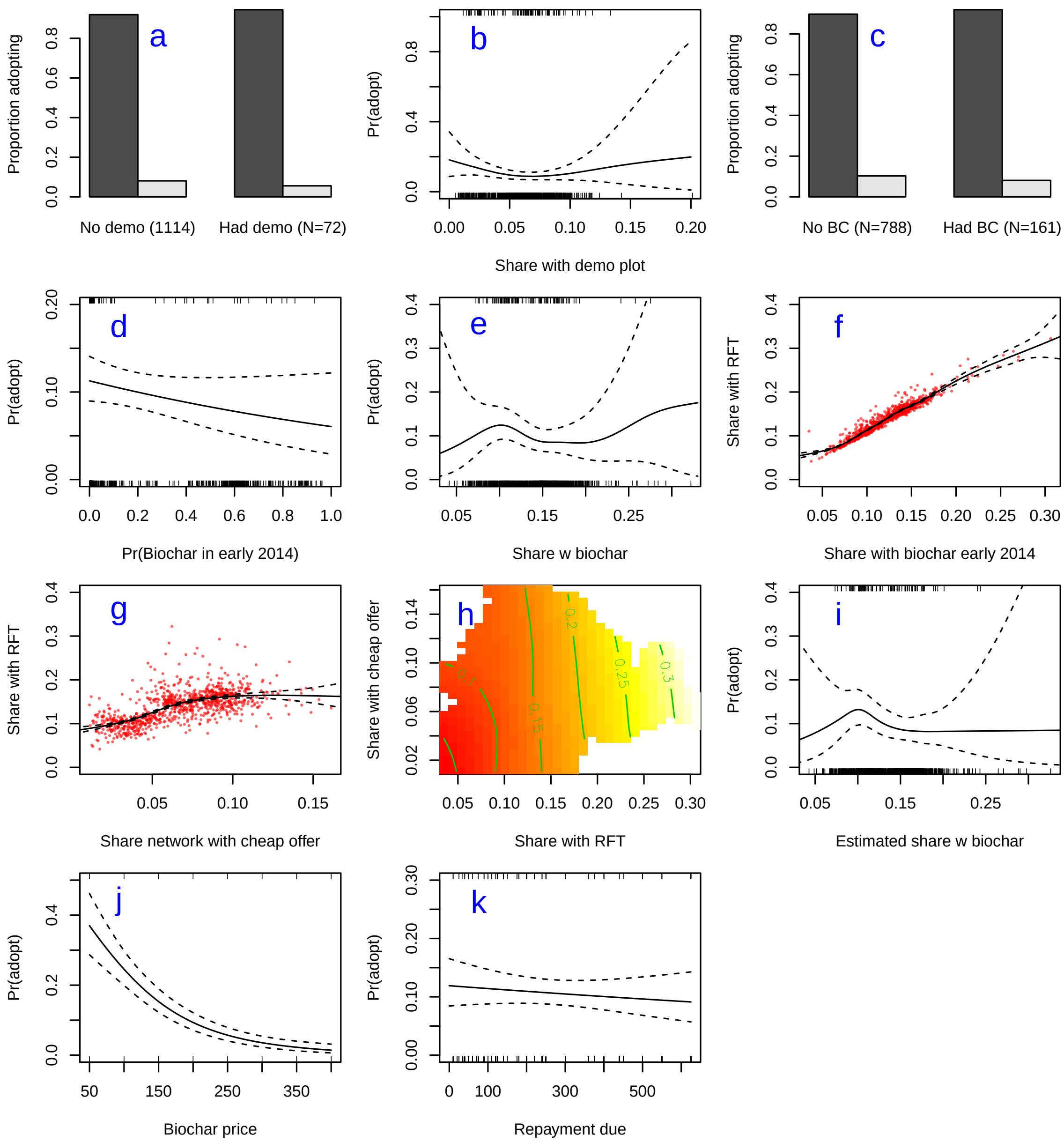


TABLE 4.5: First and second-stage results of model 4.6. Parametric terms in upper table, non-parametric terms in lower plot matrix. All coefficients and smooth function estimates given on the scale of the link function (log or logit). P-values on smooth estimates are associated with Wood's [80] reduced-rank Wald test for joint equality to zero among the coefficients comprising a smoothed term. In addition to the terms shown, all models control for income, RFT repayment price, and distance from the center of the project zone. In addition, second-stage models control for smooth functions of the residuals of the first stage models. These estimates are omitted to save space, but are available on request.

	Dependent variable			
	Adopted LR2014	Share w/ BC LR2014	Adopt SR2014	# bags SR2014
Intercept	-6.59^{***}	-1.98^{***}	-4.23^{***}	-4.11^{***}
	(1.19)	(0.01)	(0.57)	(0.55)
Mateka (zone)	-0.24	-0.09^{***}	1.91^{**}	1.84^{**}
	(0.88)	(0.01)	(0.65)	(0.62)
Mukwa (zone)	-0.45	0.00	1.07	1.26^{*}
	(0.76)	(0.01)	(0.59)	(0.56)
Nasianda (zone)	-0.05	0.02	2.56^{***}	2.42^{***}
	(0.71)	(0.01)	(0.65)	(0.62)
Gender	-0.37	0.01	-0.07	-0.04
	(0.36)	(0.00)	(0.32)	(0.29)
Heard of ACON	1.79^{**}	0.00	0.21	0.33
	(0.68)	(0.01)	(0.49)	(0.46)
Ever used char as fert.	-1.27	0.00	-0.37	0.01
	(0.86)	(0.01)	(0.62)	(0.53)
Heard of Re:char	-0.10	-0.01	0.18	-0.30
	(0.70)	(0.01)	(0.67)	(0.66)
Heard of Rutuba	0.01	0.00	1.19^{*}	0.96
	(0.74)	(0.01)	(0.59)	(0.56)
Got demo plot	2.12^{**}	-0.01	-0.94	-0.86
	(0.66)	(0.01)	(0.60)	(0.56)
Biochar early 2014			-0.86	-0.97
			(0.63)	(0.53)
family(*link*)	Binomial(*logit*)	Gamma(*log*)	Binomial(*logit*)	Neg. Bin.(*log*)
Deviance explained	0.62	0.96	0.27	0.45
Num. obs.	803	803	803	803

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

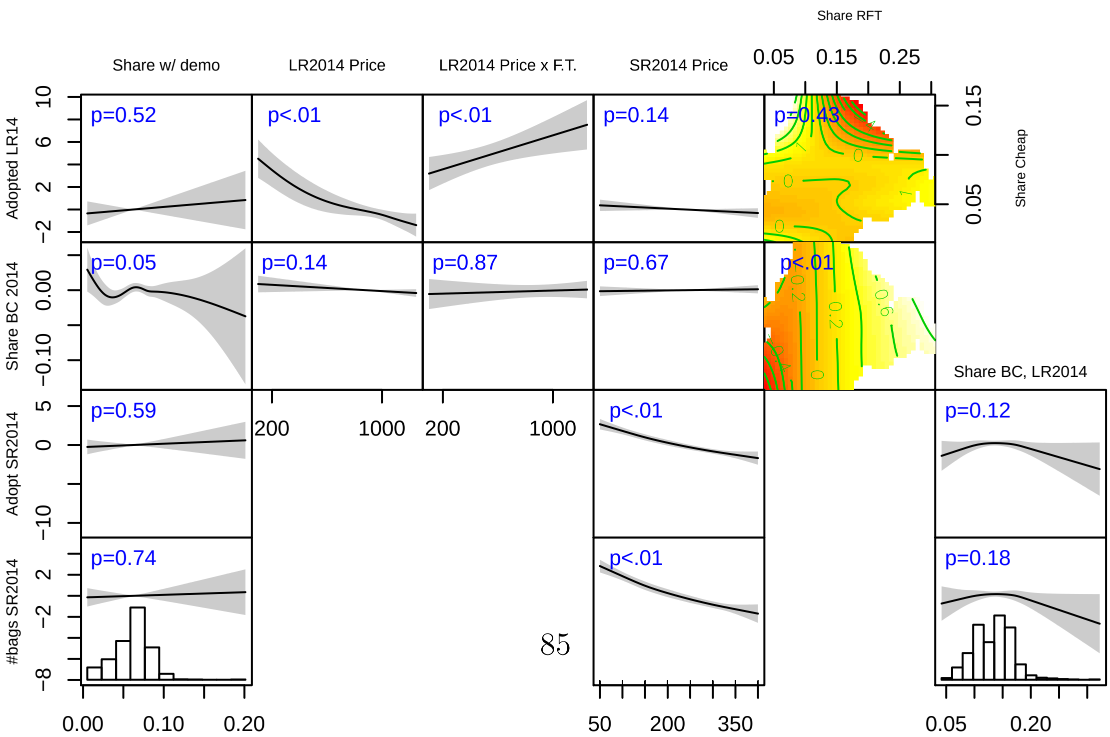


FIGURE 4.6: Prospective estimates of the welfare and implications and social return on investment of direct subsidies and provision of demonstration plots for promoting biochar. All estimates are formed from predicted values (for a median farmer) of models relating adoption to randomized treatments, as well as yield increase estimated from results of demonstration plots. All plots share x and y-axes; the percentage subsidy applied to a bag of biochar, and the number of demonstrations given – which equivalently is the proportion in this project's sample given a demonstration plot. (a) gives the expectation of the number of bags of biochar bought by a median farmer. (b) gives program cost in units of \$/farmer. The blue dashed line shows the ridge that maximizes adoption for a given program expenditure. (c) gives the expected change in profit for an individual farmer given their expected uptake of biochar at given prices and demonstration plot densities. (d) gives percentage social return on investment in subsidies and costs of demonstration plot establishment, defined as expected change in farmers profit less change in farmer and cost of subsidies. (e) gives additional social ROI (in percentage points) if sequestered carbon is valued at \$100/ton CO2e. (f) gives the carbon price required for the subsidization/demonstration scheme to break even if valued solely on the basis of sequestered carbon, solely for the region where the expected number of bags adopted is greater than 1.

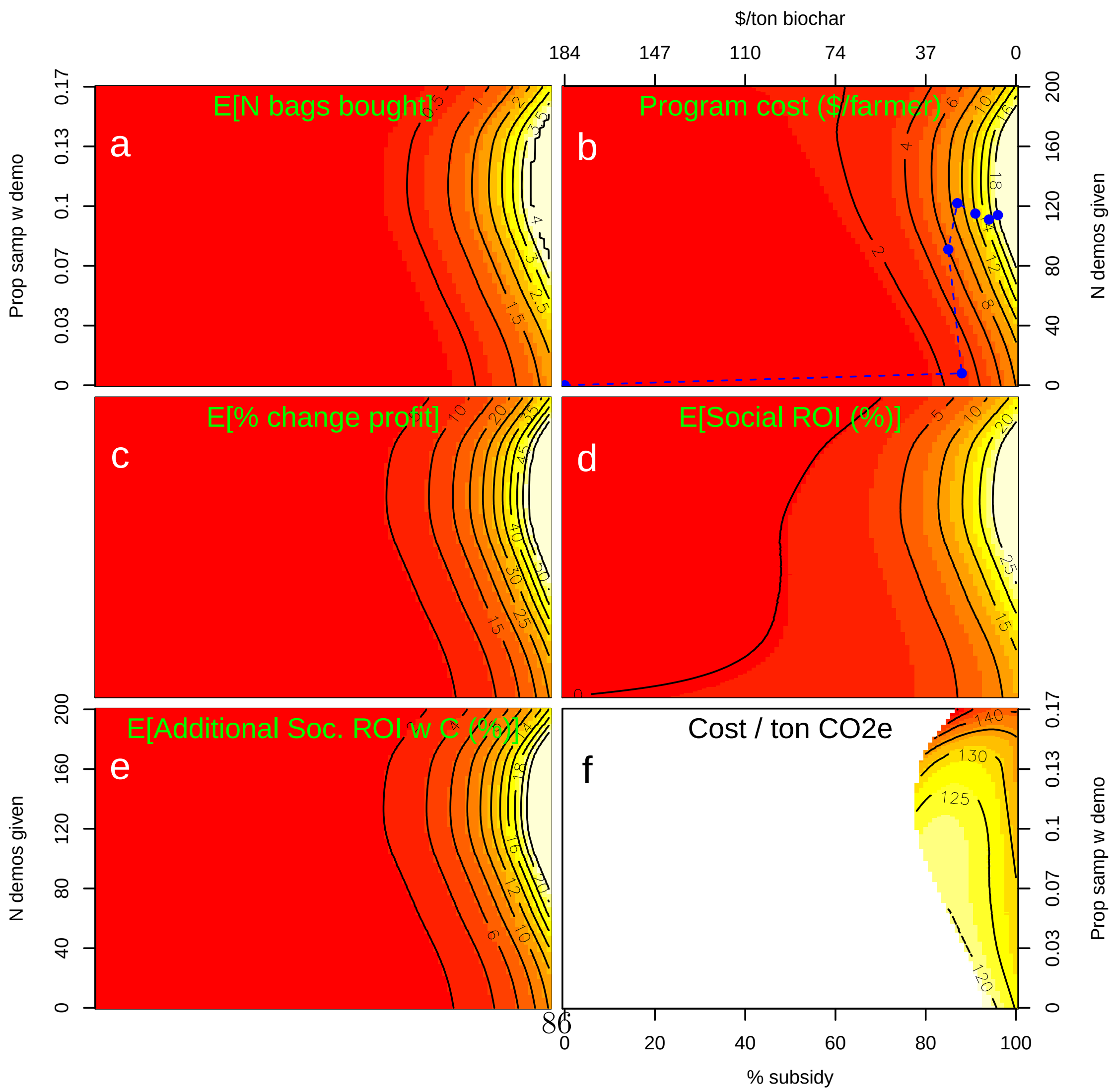
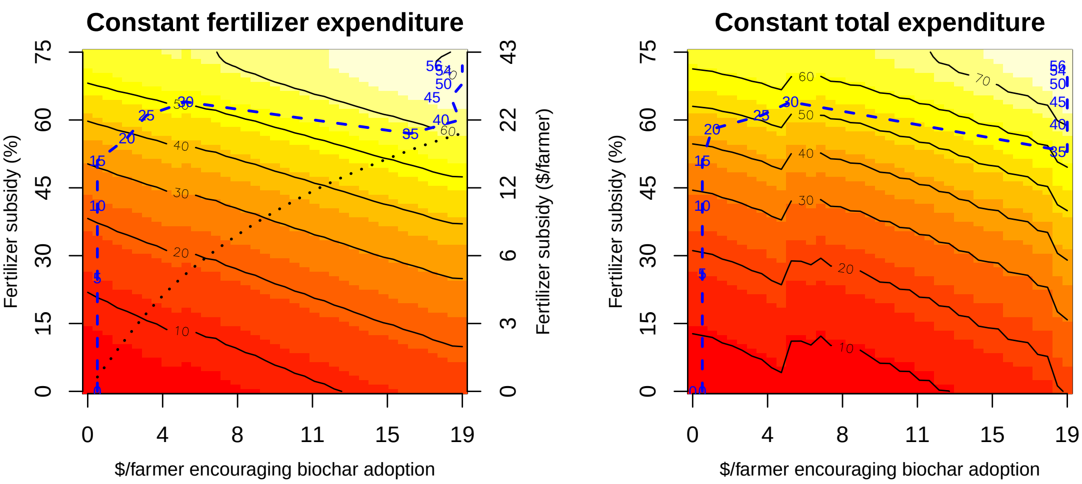


FIGURE 4.7: Percentage social return on investment for a median farmer – defined as expected farmer profit less increased farmer expenditure less program expenditure – for fertilizer subsidization (as a percentage of full cost) and biochar inducements (in units of $/farmer, along the optimal ridge given in figure 4.6b). The left figures assumes that the farmer will hold fertilizer expenditures constant as they adopt biochar, while the right figure assumes that total expenditures on biochar plus fertilizer will remain constant. The dashed blue lines give the program expenditure that maximizes social ROI; numbers in blue represent program expenditures per farmer (fertilizer and biochar inducements combined). The dotted line gives the point of equivalence between program expenditures on fertilizer and program expenditures on biochar.



Chapter 5

Biochar in Smallholder Agriculture: The Way Forward

5.1 What have we learned?

Biochar has considerable technical potential to increase crop yields in tropical soils[32]. Not only are these soils coincident with some of the world's largest "yield gaps"[29], it seems plausible that biochar might redefine attainable yields in such places, by improving various soil properties that define attainable yield. As the world enters an era in which the challenges of meeting a growing demand for food will fall upon ever-less prime agricultural land, solutions for improving their productivity are needed, and biochar appears to offer potential to meet them.

However, biochar's niche appears to be rather limited to these areas, and even to patches of rather-degraded land within these areas. In the second and third chapters of the dissertation, I found heterogeneous benefits to biochar adoption among early adopters. Yield increases tended to be higher where yields were lower, and lower where yields were higher. No yield benefit at all was found among early adopters in Vietnam. To the extent that poor soils are coincident with poverty, high potential benefit may be coincident with low effective demand. This dynamic may have been one of the explanations for the low uptake observed in the RCT described in the third paper[34]. While farmers only had one season of observations from which to learn about potential outcomes to biochar adoption, it is plausible that farmers with better potential outcomes were unable to afford to buy biochar, while those who were able to afford to buy it may have known that their own potential outcomes tended to be lower. While this is not directly testable with the data at hand, it is likely that such dynamics would become more prominent as experience with biochar increases over time.

5.2 The way forward

Heavy subsidies induced adoption by most farmers in my Kenya sample. However, for the median farmer in my Kenyan sample, public money appears to be better-spent subsidizing inorganic fertilizer than subsidizing biochar, when considering the proposition from a short-term perspective, and ignoring externalities. While this calculation might be more favorable to biochar if the sequestered carbon were valued or future benefits to crop yields were valued, it remains the case that carbon trading schemes have not reached rural Africans, and implicit discount rates are extremely high [149]. However, focus on the "typical" farmer ignores heterogeneity in response. If it were possible to know *a priori* which farmers had the best potential outcomes, then extension, education, and/or subsidies could be targeted to these farmers. With the increasing prevalence of soil testing and analysis services in Kenya and much of the rest of Africa, this prospect deserves further attention.

However, it is unclear who will provide the biochar needed for such approaches. Low effective demand among the poor will make it difficult for the private sector to do so. There may be openings for NGOs operating in the space to do so, but solutions for biochar production, along with dissemination of a bulky material to dispersed end-users will be a challenge. And while decentralized systems for biochar production may seem economical, the failure of re:char to sustain itself selling small-scale biochar production kilns suggests that behavioral issues will impede the success of this strategy. With little supply to drive learning and thereby demand, and with benefit being heterogeneous and contingent, dissemination is likely to be slow.

The problems of growing food demand in the face of environmental change are likely to worsen in the 21st century, and this does not bode well for the rural poor in sub-Saharan Africa. While biochar seems to offer one useful tool, its way forward appears impeded by economic and behavioral factors. These deserve further research. First, what can be learned about the process of technology dissemination from the experiences described in this dissertation, and how can this inform dissemination of novel technologies with a more immediately-apparent business plan? Second, what other approaches might successfully meet the same needs? Alternatively, under what conditions might biochar become economically viable? Third, who are the appropriate actors and institutions for advancing an agenda of applied research and implementation to address this problem space? The answers to these questions will be highly context dependent – and require interdisciplinary approaches – but answering them will be central to the achievement of sustainable agricultural development in an increasingly warming, crowded, and carbon-constrained future.

Bibliography

[1] UN FAO. FAOSTAT. faostat.fao.org, 2014.

[2] World Bank. *Agriculture for Development.* World Bank and Oxford University Press, 2008.

[3] BK Jack. Constraints on the adoption of agricultural technologies in developing countries. Technical report, White Paper, Agricultural Technology Adoption Initiative, J-PAL (MIT) and CEGA (UC Berkeley), 2011.

[4] P.A. Sanchez. Soil fertility and hunger in Africa. *Science*, 295(5562):2019, 2002.

[5] Nicholas Minot and Todd Benson. Fertilizer subsidies in africa: Are vouchers the answer? Technical report, International Food Policy Research Institute (IFPRI), 2009.

[6] P.P. Marenya and C.B. Barrett. State-conditional fertilizer yield response on western Kenyan farms. *American Journal of Agricultural Economics*, 91(4):991–1006, 2009.

[7] P.P. Marenya and C.B. Barrett. Soil quality and fertilizer use rates among smallholder farmers in western Kenya. *Agricultural Economics*, 40(5):561–572, 2009.

[8] S. Dercon and L. Christiaensen. Consumption risk, technology adoption and poverty traps: evidence from Ethiopia. *Journal of Development Economics*, 96(2):159–173, 2011.

[9] EMA Smaling, JJ Stoorvogel, and PN Windmeijer. Calculating soil nutrient balances in Africa at different scales. *Nutrient Cycling in Agroecosystems*, 35(3):237–250, 1993.

[10] JJ Stoorvogel, E.M.A. Smaling, and BH Janssen. Calculating soil nutrient balances in Africa at different scales. *Nutrient Cycling in Agroecosystems*, 35(3):227–235, 1993.

[11] Christopher B Barrett. Poverty traps and resource dynamics in smallholder agrarian systems. *Economics of poverty, environment and natural-resource use*, pages 17–40, 2008.

[12] Pablo Tittonell and Ken E Giller. When yield gaps are poverty traps: the paradigm of ecological intensification in african smallholder agriculture. *Field Crops Research*, 143:76–90, 2013.

[13] Jonathan A Foley, Navin Ramankutty, Kate A Brauman, Emily S Cassidy, James S Gerber, Matt Johnston, Nathaniel D Mueller, Christine O'Connell, Deepak K Ray, Paul C West, et al. Solutions for a cultivated planet. *Nature*, 478(7369):337–342, 2011.

[14] United Nations Department of Economic and Social Affairs. World Population Prospects: The 2010 Revision, 2010. http://esa.un.org/unpd/wpp/index.htm.

[15] Mark W Rosegrant, Simla Tokgoz, and Prapti Bhandary. The new normal? a tighter global agricultural supply and demand relation and its implications for food security. *American Journal of Agricultural Economics*, page aas041, 2012.

[16] W. Schlenker and D.B. Lobell. Robust negative impacts of climate change on African agriculture. *Environmental Research Letters*, 5:014010, 2010.

[17] World Bank. World Bank Data, 2014. http://data.worldbank.org/.

[18] Thomas Reardon, Christopher Barrett, Valerie Kelly, and Kimseyinga Savadogo. Policy reforms and sustainable agricultural intensification in africa. *Development policy review*, 17(4):375–395, 1999.

[19] Alain De Janvry and Elisabeth Sadoulet. Subsistence farming as a safety net for food-price shocks. *Development in Practice*, 21(4-5):472–480, 2011.

[20] J. Fairhead and M. Leach. *Misreading the African landscape: society and ecology in a forest-savanna mosaic*, volume 90. Cambridge Univ Pr, 1996.

[21] M. Tiffen, M. Mortimore, F. Gichuki, et al. *More people, less erosion: environmental recovery in Kenya.* John Wiley & Sons Ltd, 1994.

[22] Monica Van Beusekom. Negotiating development: African farmers and colonial experts at the office du niger, 1920-1960 (social history of africa). 2001.

[23] Richard P Allan and Brian J Soden. Atmospheric warming and the amplification of precipitation extremes. *Science*, 321(5895):1481–1484, 2008.

[24] Maximilian Auffhammer, Solomon M Hsiang, Wolfram Schlenker, and Adam Sobel. Using weather data and climate model output in economic analyses of climate change. *Review of Environmental Economics and Policy*, page ret016, 2013.

[25] T. Suri. Selection and comparative advantage in technology adoption. *Econometrica*, 79(1):159–209, 2011.

[26] D. Cordell, J.O. Drangert, and S. White. The story of phosphorus: Global food security and food for thought. *Global Environmental Change*, 19(2):292–305, 2009.

[27] P.A. Sanchez. *Properties and Management of Soils in the Tropics*. Wiley, 1976.

[28] N Singh, S Abiven, MS Torn, and MWI Schmidt. Fire-derived organic carbon in soil turns over on a centennial scale. *Biogeosciences*, 9:2847–2857, 2012.

[29] Nathaniel D Mueller, James S Gerber, Matt Johnston, Deepak K Ray, Navin Ramankutty, and Jonathan A Foley. Closing yield gaps through nutrient and water management. *Nature*, 490(7419):254–257, 2012.

[30] David Zilberman, Gal Hochman, Deepak Rajagopal, Steve Sexton, and Govinda Timilsina. The impact of biofuels on commodity food prices: Assessment of findings. *American Journal of Agricultural Economics*, page aas037, 2012.

[31] Stefan Dercon and Douglas Gollin. Agriculture in african development: theories and strategies. *Annu. Rev. Resour. Econ.*, 6(1):471–492, 2014.

[32] Andrew Crane-Droesch, Samuel Abiven, Simon Jeffery, and Margaret S Torn. Heterogeneous global crop yield response to biochar: a meta-regression analysis. *Environmental Research Letters*, 8(4):044049, 2013.

[33] A. Crane-Droesch, A. Claire, and A. dePinto. Biochar profitability and carbon prices: Evidence from Kenya and Vietnam, 2014. Working Paper.

[34] Andrew Crane-Droesch. Subsidies, social learning, and charcoal dust: Impact and adoption dynamics of biochar in western kenya, 2013. Working Paper.

[35] J. Lehmann and S. Joseph. *Biochar for environmental management: science and technology*. Earthscan/James & James, 2009.

[36] D. Woolf, J.E. Amonette, F.A. Street-Perrott, J. Lehmann, and S. Joseph. Sustainable biochar to mitigate global climate change. *Nature communications*, 1(5): 1–9, 2010.

[37] SP Sohi, E. Krull, E. Lopez-Capel, and R. Bol. A review of biochar and its use and function in soil. *Advances in Agronomy*, 105:47–82, 2010.

[38] K. Hammes, Margaret Torn, Andrei Lapenas, and Michael Schmidt. Centennial black carbon turnover observed in a Russian steppe soil. *Biogeosciences*, 5:1339–1350, 2008.

[39] W.G. Sombroek, F.O. Nachtergaele, and A. Hebel. Amounts, dynamics and sequestering of carbon in tropical and subtropical soils. *Ambio*, 22(7):417–426, 1993.

[40] JG Shepherd. *Geoengineering the climate: science, governance and uncertainty*. Royal Society, 2009.

[41] A. Zhang, L. Cui, G. Pan, L. Li, Q. Hussain, X. Zhang, J. Zheng, and D. Crowley. Effect of biochar amendment on yield and methane and nitrous oxide emissions from a rice paddy from Tai Lake plain, China. *Agriculture, Ecosystems & Environment*, 2010.

[42] Bhupinder Pal Singh, Blake J Hatton, Balwant Singh, Annette L Cowie, and Amrit Kathuria. Influence of biochars on nitrous oxide emission and nitrogen leaching from two contrasting soils. *Journal of environmental quality*, 39(4):1224–1235, 2010.

[43] L. Van Zwieten, B. Singh, S. Joseph, S. Kimber, A. Cowie, and Y. Chan. Biochar and emissions of non-CO2 greenhouse gases from soil. In *Biochar for Environmental Management–Science and Technology*, pages 227–249. Earthscan, 2009.

[44] S. Jeffery, F.G.A. Verheijen, M. van der Velde, and A.C. Bastos. A quantitative review of the effects of biochar application to soils on crop productivity using meta-analysis. *Agriculture, Ecosystems & Environment*, 144(1):175–187, November 2011. ISSN 0167-8809. doi: 10.1016/j.agee.2011.08.015. URL http://www.sciencedirect.com/science/article/pii/S0167880911003197.

[45] L.A. Biederman and W.S. Harpole. Biochar and its effects on plant productivity and nutrient cycling: a meta-analysis. *GCB Bioenergy*, 2012.

[46] C.J. Atkinson, J.D. Fitzgerald, and N.A. Hipps. Potential mechanisms for achieving agricultural benefits from biochar application to temperate soils: a review. *Plant and Soil*, 337(1):1–18, 2010.

[47] K.A. Spokas, K.B. Cantrell, J.M. Novak, D.W. Archer, J.A. Ippolito, H.P. Collins, A.A. Boateng, I.M. Lima, M.C. Lamb, A.J. McAloon, et al. Biochar: A synthesis of its agronomic impact beyond carbon sequestration. *Journal of Environmental Quality*, 41(4):973–989, 2012.

[48] L. Van Zwieten, S. Kimber, S. Morris, KY Chan, A. Downie, J. Rust, S. Joseph, and A. Cowie. Effects of biochar from slow pyrolysis of papermill waste on agronomic performance and soil fertility. *Plant and soil*, 327(1):235–246, 2010.

[49] C. Steiner, W.G. Teixeira, J. Lehmann, T. Nehls, J.L.V. de Macêdo, W.E.H. Blum, and W. Zech. Long term effects of manure, charcoal and mineral fertilization on crop production and fertility on a highly weathered Central Amazonian upland soil. *Plant and Soil*, 291(1):275–290, 2007.

[50] Gerard Cornelissen, Vegard Martinsen, Victor Shitumbanuma, Vanja Alling, Gijs D Breedveld, David W Rutherford, Magnus Sparrevik, Sarah E Hale, Alfred Obia, and Jan Mulder. Biochar effect on maize yield and soil characteristics in five conservation farming sites in zambia. *Agronomy*, 3(2):256–274, 2013.

[51] D. Rubin. *Multiple Imuutation for Nonresponse in Surveys*. Wiley Online Library, 1987.

[52] S.N Wood. *Generalized Additive Models: An Introduction with R*. Chapman and Hall/CRC, 2006.

[53] S. N. Wood. Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *Journal of the Royal Statistical Society (B)*, 73(1):3–36, 2011.

[54] Simon G Thompson and Julian Higgins. How should meta-regression analyses be undertaken and interpreted? *Statistics in medicine*, 21(11):1559–1573, 2002.

[55] NH Batjes. Isric-wise global data set of derived soil properties on a 0.5 by 0.5 degree grid (version 3.0). Technical report, Wageningen: ISRIC-World Soil Information, 2005.

[56] J. Gurevitch, P.S. Curtis, and M.H. Jones. Meta-analysis in ecology. *Advances in Ecological Research*, 32:199–247, 2001.

[57] L.V. Hedges, J. Gurevitch, and P.S. Curtis. The meta-analysis of response ratios in experimental ecology. *Ecology*, 80(4):1150–1156, 1999.

[58] M. Keiluweit, P.S. Nico, M.G. Johnson, and M. Kleber. Dynamic molecular structure of plant biomass-derived black carbon (biochar). *Environmental science & technology*, 44(4):1247–1253, 2010.

[59] J.H. Yuan, R.K. Xu, and H. Zhang. The forms of alkalis in the biochar produced from crop residues at different temperatures. *Bioresource technology*, 2010.

[60] W. Viechtbauer. Bias and efficiency of meta-analytic variance estimators in the random-effects model. *Journal of Educational and Behavioral Statistics*, 30(3):261–293, 2005.

[61] C.S. Berkey, D.C. Hoaglin, F. Mosteller, and GA Colditz. A random-effects regression model for meta-analysis. *Statistics in medicine*, 14(4):395–411, 2007.

[62] Bethany A Bell, John M Ferron, and Jeffrey D Kromrey. Cluster size in multilevel models: The impact of sparse data structures on point and interval estimates in two-level models. In *JSM Proceedings*, pages 1122–1129. American Statistical Association Denver, CO, 2008.

[63] R.J.A. Little and D.B. Rubin. *Statistical analysis with missing data*, volume 333. Wiley-Interscience, 2002.

[64] I.R. White, P. Royston, and A.M. Wood. Multiple imputation using chained equations: Issues and guidance for practice. *Statistics in medicine*, 30(4):377–399, 2011.

[65] A. Gelman and J. Hill. *Data analysis using regression and multilevel/hierarchical models*. Cambridge University Press, 2006.

[66] Trevor Hastie and Robert Tibshirani. Generalized additive models. *Statistical science*, pages 297–310, 1986.

[67] N. Ramankutty. Global cropland and pasture data: 1700-2007, 2011.

[68] DL Jones, J. Rousk, G. Edwards-Jones, TH DeLuca, and DV Murphy. Biochar-mediated changes in soil quality and plant growth in a three year field trial. *Soil Biology and Biochemistry*, 2011.

[69] J. Major, M. Rondon, D. Molina, S.J. Riha, and J. Lehmann. Maize yield and nutrition during 4 years after biochar application to a Colombian savanna oxisol. *Plant and soil*, 333(1):117–128, 2010.

[70] R.S. Quilliam, K.A. Marsden, C. Gertler, J. Rousk, T.H. DeLuca, and D.L. Jones. Nutrient dynamics, microbial growth and weed emergence in biochar amended soil are influenced by time since application and reapplication rate. *Agriculture, Ecosystems & Environment*, 158:192–199, 2012.

[71] SM Haefele, Y. Konboon, W. Wongboon, S. Amarante, AA Maarifat, EM Pfeiffer, and C. Knoblauch. Effects and fate of biochar from rice residues in rice-based systems. *Field Crops Research*, 2011.

[72] Emma C Suddick and Johan Six. An estimation of annual nitrous oxide emissions and soil quality following the amendment of high temperature walnut shell biochar and compost to a small scale vegetable crop rotation. *Science of The Total Environment*, 2013.

[73] C. Rumpel, V. Chaplot, O. Planchon, J. Bernadou, C. Valentin, and A. Mariotti. Preferential erosion of black carbon on steep slopes with slash and burn agriculture. *Catena*, 65(1):30–40, 2006.

[74] Rudolf Jaffé, Yan Ding, Jutta Niggemann, Anssi V Vähätalo, Aron Stubbins, Robert GM Spencer, John Campbell, and Thorsten Dittmar. Global charcoal mobilization from soils via dissolution and riverine transport to the oceans. *Science*, 340(6130):345–347, 2013.

[75] J. Major, J. Lehmann, M. Rondon, and C. Goodale. Fate of soil-applied black carbon: downward migration, leaching and soil respiration. *Global Change Biology*, 16(4):1366–1379, 2010.

[76] J.E. Amonette, S. Joseph, J. Lehmann, and S. Joseph. Characteristics of Biochar-Micro-chemical Properties. In *Biochar for Environmental Management: Science and Technology*. Earthscan, 2009.

[77] S. Brodowski, B. John, H. Flessa, and W. Amelung. Aggregate-occluded black carbon in soil. *European Journal of Soil Science*, 57(4):539–546, 2006.

[78] M.W.I. Schmidt, M.S. Torn, S. Abiven, T. Dittmar, G. Guggenberger, I.A. Janssens, M. Kleber, I. Kögel-Knabner, J. Lehmann, D.A.C. Manning, et al. Persistence of soil organic matter as an ecosystem property. *Nature*, 478(7367):49–56, 2011.

[79] World Bank. Carbon Sequestration in Agricultural Soils. Technical report, World Bank, 2012.

[80] Simon N Wood. On p-values for smooth components of an extended generalized additive model. *Biometrika*, 100(1):221–228, 2013.

[81] T.G. Vågen, R. Lal, and BR Singh. Soil carbon sequestration in sub-Saharan Africa: a review. *Land Degradation & Development*, 16(1):53–71, 2005.

[82] Sebastian B Scholz and Thomas Sembres. *Biochar Systems for Smallholders in Developing Countries: Leveraging Current Knowledge and Exploring Future Potential for Climate-Smart Agriculture*. World Bank Publications, 2014.

[83] A.R. Zimmerman. Abiotic and microbial oxidation of laboratory-produced black carbon (biochar). *Environmental science & technology*, 44(4):1295–1301, 2010.

[84] D. Solomon, J. Lehmann, J. Kinyangi, W. Amelung, I. Lobe, A. Pell, S. Riha, S. Ngoze, L. Verchot, D. Mbugua, et al. Long-term impacts of anthropogenic perturbations on dynamics and speciation of organic carbon in tropical forest and subtropical grassland ecosystems. *Global Change Biology*, 13(2):511–530, 2007.

[85] B. Glaser and W.I. Woods. *Amazonian Dark Earths: Explorations in space and time*. Springer Verlag, 2004.

[86] Stefan Jirka and Thayer Tomlinson. 2013 state of the biochar industry. 2013.

[87] D. Laird, P. Fleming, B. Wang, R. Horton, and D. Karlen. Biochar impact on nutrient leaching from a Midwestern agricultural soil. *Geoderma*, 158(3-4):436–442, 2010.

[88] J. Lehmann, J. Pereira da Silva, C. Steiner, T. Nehls, W. Zech, and B. Glaser. Nutrient availability and leaching in an archaeological Anthrosol and a Ferralsol of the Central Amazon basin: fertilizer, manure and charcoal amendments. *Plant and Soil*, 249(2):343–357, 2003.

[89] J. Major, C. Steiner, A. Downie, J. Lehmann, and S. Joseph. Biochar effects on nutrient leaching. In *Biochar for Environmental Management: Science and Technology*, pages 271–284. Earthscan, 2009.

[90] J. Major, M. Rondon, D. Molina, S.J. Riha, and J. Lehmann. Nutrient Leaching in a Colombian Savanna Oxisol Amended with Biochar. *Journal of Environmental Quality*, 41, 2011.

[91] A. Zhang, Y. Liu, G. Pan, Q. Hussain, L. Li, J. Zheng, and X. Zhang. Effect of biochar amendment on maize yield and greenhouse gas emissions from a soil organic carbon poor calcareous loamy soil from central china plain. *Plant and soil*, 351(1): 263–275, 2012.

[92] B.A. McCarl, C. Peacocke, R. Chrisman, CC Kung, and RD Sands. Economics of biochar production, utilisation and GHG offsets. In *Biochar for Environmental Management: Science and Technology*. 2009.

[93] Suzette P Galinato, Jonathan K Yoder, and David Granatstein. The economic value of biochar in crop production and carbon sequestration. *Energy Policy*, 39 (10):6344–6350, 2011.

[94] Kelli G Roberts, Brent A Gloy, Stephen Joseph, Norman R Scott, and Johannes Lehmann. Life cycle assessment of biochar systems: Estimating the energetic, economic, and climate change potential. *Environmental Science & Technology*, 44 (2):827–833, 2009.

[95] Morgan M Williams and James C Arnott. A comparison of variable economic costs associated with two proposed biochar application methods. *Annals of Environmental Science*, 4(1):4, 2010.

[96] Abbie Clare, Andrew Barnes, John McDonagh, and Simon Shackley. From rhetoric to reality: farmer perspectives on the economic potential of biochar in china. *International Journal of Agricultural Sustainability*, 2014. doi: 10.1080/14735903.2014.927711.

[97] Jonathan Yoder, Suzette Galinato, David Granatstein, and Manuel Garcia-Pérez. Economic tradeoff between biochar and bio-oil production via pyrolysis. *biomass and bioenergy*, 35(5):1851–1862, 2011.

[98] P. Blackwell, E. Krull, G. Butler, A. Herbert, and Z. Solaiman. Effect of banded biochar on dryland wheat production and fertiliser use in south-western Australia: an agronomic and economic perspective. *Australian Journal of Soil Research*, 48 (7):531–545, 2010.

[99] K. Pratt and D. Moran. Evaluating the cost-effectiveness of global biochar mitigation potential. *Biomass and Bioenergy*, 34(8):1149–1158, 2010.

[100] Pietro Galgani, Ester van der Voet, and Gijsbert Korevaar. Composting, anaerobic digestion and biochar production in ghana. environmental–economic assessment in the context of voluntary carbon markets. *Waste Management*, 2014.

[101] J. Pender, S. Ehui, et al. *Strategies for sustainable land management in the East African highlands*. Intl Food Policy Research Inst, 2006.

[102] Nguyen Huu Dung, Tran Chi Thien, NV Hong, Nguyen Thi Loc, DV Minh, Trinh Dinh Thau, HTL Nguyen, Nguyen Tan Phong, and Thai Thanh Son. *Impact of agro-chemical use on productivity and health in Vietnam*. Economy and environment program for Southeast Asia (EEPSEA), 1999.

[103] Jianguo Liu and Jared Diamond. China's environment in a globalizing world. *Nature*, 435(7046):1179–1186, 2005.

[104] XT Ju, CL Kou, P Christie, ZX Dou, and FS Zhang. Changes in the soil environment from excessive application of fertilizers and manures to two contrasting intensive cropping systems on the north china plain. *Environmental Pollution*, 145 (2):497–506, 2007.

[105] Nguyen Huu Huan, LV Thiet, HV Chien, and KL Heong. Farmers' participatory evaluation of reducing pesticides, fertilizers and seed rates in rice farming in the mekong delta, vietnam. *Crop Protection*, 24(5):457–464, 2005.

[106] Xiao-Tang Ju, Guang-Xi Xing, Xin-Ping Chen, Shao-Lin Zhang, Li-Juan Zhang, Xue-Jun Liu, Zhen-Ling Cui, Bin Yin, Peter Christie, Zhao-Liang Zhu, et al. Reducing environmental risk by improving n management in intensive chinese agricultural systems. *Proceedings of the National Academy of Sciences*, 106(9):3041–3046, 2009.

[107] Jikun Huang, Ruifa Hu, Jianmin Cao, and Scott Rozelle. Training programs and in-the-field guidance to reduce china's overuse of fertilizer without hurting profitability. *jOurnal Of SOil and water cOnServatiOn*, 63(5):165A–167A, 2008.

[108] G Yohe, R Lasco, QK Ahmad, S Cohen, T Janetos, R Perez, K Ebi, P Romero Lankao, E Malone, T Malone, et al. Perspectives on climate change and sustainability 3. *change*, 25(48):49, 2007.

[109] Francesca de Nicola and Xavier Giné. How accurate are recall data? evidence from coastal india. *Journal of Development Economics*, 106:52–65, 2014.

[110] David A Harville. Confidence intervals and sets for linear combinations of fixed and random effects. *Biometrics*, pages 403–407, 1976.

[111] David A Harville. Maximum likelihood approaches to variance component estimation and to related problems. *Journal of the American Statistical Association*, 72 (358):320–338, 1977.

[112] Yunying Fang, Bhupinder Pal Singh, and Balwant Singh. Temperature sensitivity of biochar and native carbon mineralisation in biochar-amended soils. *Agriculture, Ecosystems & Environment*, 2014.

[113] D. Torres-Rojas, J. Lehmann, P. Hobbs, S. Joseph, and H. Neufeldt. Biomass availability, energy consumption and biochar production in rural households of Western Kenya. *Biomass and Bioenergy*, 2011.

[114] Sebastian Meyer, Bruno Glaser, and Peter Quicker. Technical, economical, and climate-related aspects of biochar production technologies: a literature review. *Environmental science & technology*, 45(22):9473–9483, 2011.

[115] R Brown. Biochar production technology. In *Biochar for Environmental Management: Science and Technology*, pages 127–146. Earthscan, London, 2009.

[116] Kyle Crombie and Ondřej Mašek. Pyrolysis biochar systems, balance between bioenergy and carbon sequestration. *GCB Bioenergy*, 2014.

[117] Thea Whitman, Charles F Nicholson, Dorisel Torres, and Johannes Lehmann. Climate change impact of biochar cook stoves in western kenyan farm households: system dynamics model analysis. *Environmental science & technology*, 45(8):3687–3694, 2011.

[118] Magnus Sparrevik, John L Field, Vegard Martinsen, Gijs D Breedveld, and Gerard Cornelissen. Life cycle assessment to evaluate the environmental impact of biochar implementation in conservation agriculture in zambia. *Environmental science & technology*, 47(3):1206–1215, 2013.

[119] James Jetter, Yongxin Zhao, Kirk R Smith, Bernine Khan, Tiffany Yelverton, Peter DeCarlo, and Michael D Hays. Pollutant emissions and energy efficiency under controlled conditions for household biomass cookstoves and implications for metrics useful in setting international test standards. *Environmental science & technology*, 46(19):10827–10834, 2012.

[120] I. Ruiz-Mercado, O. Masera, H. Zamora, and K.R. Smith. Adoption and sustained use of improved cookstoves. *Energy Policy*, 2011.

[121] Jason Burwen and David I Levine. A rapid assessment randomized-controlled trial of improved cookstoves in rural ghana. *Energy for Sustainable Development*, 16(3): 328–338, 2012.

[122] Rema Hanna, Esther Duflo, and Michael Greenstone. Up in smoke: the influence of household behavior on the long-run impact of improved cooking stoves. Technical report, National Bureau of Economic Research, 2012.

[123] David I Levine, Theresa Beltramo, Garrick Blalock, and Carolyn Cotterman. What impedes efficient adoption of products? evidence from randomized variation in sales offers for improved cookstoves in uganda. 2013.

[124] Marco A Quiñones, Norman E Borlaug, and Christopher R Dowswell. A fertilizer-based green revolution for africa. *Replenishing soil fertility in Africa*, (replenishingsoi):81–95, 1997.

[125] Charles Mann. Reseeding the green revolution. *Science*, 277(5329):1038–1043, 1997.

[126] Gary Toenniessen, Akinwumi Adesina, and Joseph DeVries. Building an alliance for a green revolution in africa. *Annals of the New York Academy of Sciences*, 1136 (1):233–242, 2008.

[127] Norman E Borlaug. The green revolution revisited and the road ahead. *Special 30th Anniversary Lecture, Norwegian Nobel Institute, Oslo*, 2000.

[128] J.M. Kimetu, J. Lehmann, S.O. Ngoze, D.N. Mugendi, J.M. Kinyangi, S. Riha, L. Verchot, J.W. Recha, and A.N. Pell. Reversibility of soil productivity decline with organic matter of differing quality along a degradation gradient. *Ecosystems*, 11(5):726–739, 2008.

[129] Giacomo Branca, Timm Tennigkeit, Wendy Mann, and Leslie Lipper. Identifying opportunities for climate-smart agriculture investments in africa. *FAO, Rome*, 2011.

[130] Rattan Lal. Soil carbon sequestration to mitigate climate change. *Geoderma*, 123 (1):1–22, 2004.

[131] R. Lal. Enhancing crop yields in the developing countries through restoration of the soil organic carbon pool in agricultural lands. *Land Degradation & Development*, 17(2):197–209, 2006.

[132] D.E. Mercer. Adoption of agroforestry innovations in the tropics: A review. *Agroforestry systems*, 61(1):311–328, 2004.

[133] O.C. Ajayi, F.K. Akinnifesi, G. Sileshi, and S. Chakeredza. Adoption of renewable soil fertility replenishment technologies in the southern African region: Lessons learnt and the way forward. *Natural Resources Forum*, 31(4):306–317, 2007.

[134] Leslie Lipper, B Neves, A Wilkes, T Tennigkeit, P Gerber, B Henderson, G Branca, W Mann, et al. *Climate change mitigation finance for smallholder agriculture: a guide book to harvesting soil carbon sequestration benefits.* Food and Agriculture Organization of the United Nations (FAO), 2011.

[135] Timothy Pearson, Sarah Walker, and Sandra Brown. Sourcebook for land use, land-use change and forestry projects. 2013.

[136] John M Antle and Bocar Diagana. Creating incentives for the adoption of sustainable agricultural practices in developing countries: the role of soil carbon sequestration. *American Journal of Agricultural Economics*, 85(5):1178–1184, 2003.

[137] Michael Jerry Antal and Morten Grønli. The art, science, and technology of charcoal production. *Industrial & Engineering Chemistry Research*, 42(8):1619–1640, 2003.

[138] M.P.W. Schneider, M. Hilf, U.F. Vogt, and M.W.I. Schmidt. The benzene polycarboxylic acid (BPCA) pattern of wood pyrolyzed between 200C and 1000C. *Organic Geochemistry*, 2010.

[139] ES Krull, JA Baldock, JO Skjemstad, and RJ Smernik. In *Characteristics of biochar: organo-chemical properties. In 'Biochar for environmental management,.(Eds J Lehmann, S Joseph) pp. 53-65.* Earthscan Publications Ltd: London, 2009.

[140] WG Sombroek. *Amazon soils.* Landbouwhogeschool, 1966.

[141] B. Glaser, E. Balashov, L. Haumaier, G. Guggenberger, and W. Zech. Black carbon in density fractions of anthropogenic soils of the Brazilian Amazon region. *Organic Geochemistry*, 31(7-8):669–678, 2000.

[142] B. Glaser, L. Haumaier, G. Guggenberger, and W. Zech. The'Terra Preta' phenomenon: a model for sustainable agriculture in the humid tropics. *Naturwissenschaften*, 88(1):37–41, 2001.

[143] E. Duflo, M. Kremer, and J. Robinson. How high are rates of return to fertilizer? Evidence from field experiments in Kenya. *The American Economic Review*, 98(2): 482–488, 2008.

[144] E.M. Rogers. *Diffusion of innovation*, volume 4. New York: Free Press, 1962.

[145] Z. Griliches. Hybrid corn: An exploration in the economics of technological change. *Econometrica, Journal of the Econometric Society*, pages 501–522, 1957.

[146] Michael R Carter, Rachid Laajaj, and Dean Yang. Subsidies and the persistence of technology adoption: Field experimental evidence from mozambique. 2014.

[147] P. Dupas. Short-run subsidies and long-run adoption of new health products: Evidence from a field experiment. Technical report, National Bureau of Economic Research, 2012.

[148] Uma J Lele, Robert E Christiansen, and Kundhavi Kadiresan. Fertilizer policy in africa: lessons from development programs and adjustment lending, 1970-87. *MADIA discussion paper (USA)*, 1989.

[149] E. Duflo, M. Kremer, and J. Robinson. Nudging farmers to use fertilizer: theory and experimental evidence from Kenya, 2009.

[150] Glenn Denning, Patrick Kabambe, Pedro Sanchez, Alia Malik, Rafael Flor, Rebbie Harawa, Phelire Nkhoma, Colleen Zamba, Clement Banda, Chrispin Magombo, et al. Input subsidies to improve smallholder maize productivity in malawi: Toward an african green revolution. *PLoS biology*, 7(1):e1000023, 2009.

[151] Charles F Manski. Identification of endogenous social effects: The reflection problem. *The review of economic studies*, 60(3):531–542, 1993.

[152] Nicholas Magnan, David J Spielman, Travis J Lybbert, and Kajal Gulati. Leveling with friends: Social networks and indian farmers' demand for agricultural custom hire services. In *Selected paper for the 2013 American Applied Economics Association annual meeting, August*, volume 5, 2013.

[153] O. Bandiera and I. Rasul. Social networks and technology adoption in northern mozambique*. *The Economic Journal*, 116(514):869–902, 2006.

[154] Jing Cai, Alain De Janvry, and Elisabeth Sadoulet. Social networks and the decision to insure. 2013.

[155] Pramila Krishnan and Manasa Patnam. Neighbors and extension agents in ethiopia: Who matters more for technology adoption? *American Journal of Agricultural Economics*, 96(1):308–327, 2014.

[156] K. Munshi. Social learning in a heterogeneous population: technology diffusion in the Indian Green Revolution. *Journal of Development Economics*, 73(1):185–213, 2004.

[157] Annemie Maertens. Who cares what others think (or do)? social learning, social pressures and imitation in cotton farming in india. *Unpublished, University of Pittsburgh, Pittsburg, PA, US*, 2012.

[158] T.G. Conley and C.R. Udry. Learning about a new technology: Pineapple in Ghana. *The American Economic Review*, 100(1):35–69, 2010.

[159] Margarita Genius, Phoebe Koundouri, Céline Nauges, and Vangelis Tzouvelekas. Information transmission in irrigation technology adoption and diffusion: social learning, extension services, and spatial effects. *American Journal of Agricultural Economics*, 96(1):328–344, 2014.

[160] Abhijit V Banerjee. A simple model of herd behavior. *The Quarterly Journal of Economics*, pages 797–817, 1992.

[161] Wim G Sombroek, HMH Braun, BJA Van der Pouw, et al. *Exploratory soil map and agro-climatic zone map of Kenya, 1980. Scale 1: 1,000,000.* Kenya Soil Survey, 1982.

[162] S.N. Wood. Low-rank scale-invariant tensor product smooths for generalized additive mixed models. *Biometrics*, 62(4):1025–1036, 2006.

[163] D. Ruppert, M.P. Wand, and R.J. Carroll. *Semiparametric regression*, volume 12. Cambridge University Press, 2003.

[164] Giampiero Marra and Rosalba Radice. A flexible instrumental variable approach. *Statistical Modelling*, 11(6):581–603, 2011.

[165] James J Heckman. Dummy endogenous variables in a simultaneous equation system. *Econometrica: Journal of the Econometric Society*, pages 931–959, 1978.

[166] Whitney K Newey and James L Powell. Instrumental variable estimation of nonparametric models. *Econometrica*, 71(5):1565–1578, 2003.

[167] A.G. Chandrasekhar and R. Lewis. Econometrics of sampled networks. Technical report, MIT working paper, 2011.

[168] Atanu Saha. Risk preference estimation in the nonlinear mean standard deviation approach. *Economic Inquiry*, 35(4):770–782, 1997.

[169] G. Feder, R.E. Just, and D. Zilberman. Adoption of agricultural innovations in developing countries: A survey. *Economic development and cultural change*, pages 255–298, 1985.

[170] M. Kremer and E. Miguel. The illusion of sustainability. *Quarterly Journal of Economics*, 122:1007–65, 2007.

[171] Dean Karlan and Jonathan Zinman. Lying about borrowing. *Journal of the European Economic Association*, 6(2-3):510–521, 2008.

[172] Marianne Bertrand and Sendhil Mullainathan. Do people mean what they say? implications for subjective survey data. *American Economic Review*, pages 67–72, 2001.

[173] T.E. Bodner. What improves with increased missing data imputations? *Structural Equation Modeling*, 15(4):651–675, 2008.

[174] J.W. Graham, A.E. Olchowski, and T.D. Gilreath. How many imputations are really needed? some practical clarifications of multiple imputation theory. *Prevention Science*, 8(3):206–213, 2007.

[175] P.T. Luyen, D.N. Khang, and TR Preston. Effects of biochar from gasifier stove and effluent from biodigester on growth of maize in acid and fertile soils. *Livestock Research for Rural Development*, 24(5), 2012.

[176] X. Peng, LL Ye, CH Wang, H. Zhou, and B. Sun. Temperature-and duration-dependent rice straw-derived biochar: Characteristics and its effects on soil properties of an Ultisol in southern China. *Soil and Tillage Research*, 2011.

[177] L. Rodriguez and T.R. Preston. Effect of biochar and biodigester effluent on growth of maize in acid soils. *Livestock Research for Rural Development*, 30(1), 2009.

[178] D. Noguera, M. Rondón, K.R. Laossi, V. Hoyos, P. Lavelle, M.H. Cruz de Carvalho, and S. Barot. Contrasted effect of biochar and earthworms on rice growth and resource allocation in different soils. *Soil biology and Biochemistry*, 42(7):1017–1027, 2010.

[179] KY Chan, L. Van Zwieten, I. Meszaros, A. Downie, and S. Joseph. Using poultry litter biochars as soil amendments. *Australian Journal of Soil Research*, 46(5): 437–444, 2008.

[180] KY Chan, L. Van Zwieten, I. Meszaros, A. Downie, and S. Joseph. Agronomic values of greenwaste biochar as a soil amendment. *Australian Journal of Soil Research*, 45 (8):629–634, 2007.

[181] C.H. Cheng, J. Lehmann, and M.H. Engelhard. Natural oxidation of black carbon in soils: changes in molecular form and surface charge along a climosequence. *Geochimica et Cosmochimica Acta*, 72(6):1598–1610, 2008.

[182] SD Joseph, M. Camps-Arbestain, Y. Lin, P. Munroe, CH Chia, J. Hook, L. Van Zwieten, S. Kimber, A. Cowie, BP Singh, et al. An investigation into the reactions of biochar in soil. *Australian Journal of Soil Research*, 48(7):501–515, 2010.

[183] J.F. Artiola, C. Rasmussen, and R. Freitas. Effects of a biochar-amended alkaline soil on the growth of romaine lettuce and bermudagrass. *Soil Science*, 177(9): 561–570, 2012.

[184] J.W. Gaskin, R.A. Speir, K. Harris, KC Das, R.D. Lee, L.A. Morris, and D.S. Fisher. Effect of Peanut Hull and Pine Chip Biochar on Soil Nutrients, Corn Nutrient Status, and Yield. *Agronomy Journal*, 102:623–633, 2010.

[185] D. Güereña, J. Lehmann, K. Hanley, A. Enders, C. Hyland, and S. Riha. Nitrogen dynamics following field application of biochar in a temperate north american maize-based production system. *Plant and Soil*, pages 1–16, 2012.

[186] H. Schulz and B. Glaser. Effects of biochar compared to organic and inorganic fertilizers on soil quality and plant growth in a greenhouse experiment. *Journal of Plant Nutrition and Soil Science*, 2012.

[187] FP Vaccari, S. Baronti, E. Lugato, L. Genesio, S. Castaldi, F. Fornasier, and F. Miglietta. Biochar as a strategy to sequester carbon and increase yield in durum wheat. *European Journal of Agronomy*, 2011.

[188] A. Zhang, R. Bian, G. Pan, L. Cui, Q. Hussain, L. Li, J. Zheng, J. Zheng, X. Zhang, X. Han, et al. Effects of biochar amendment on soil quality, crop yield and greenhouse gas emission in a chinese rice paddy: a field study of 2 consecutive rice growing cycles. *Field Crops Research*, 127:153–160, 2012.

[189] W.H.U. Sukartono, Z. Kusuma, and W.H. Nugroho. Soil fertility status, nutrient uptake, and maize (zea mays l.) yield following biochar and cattle manure application on sandy soils of lombok, indonesia. *Journal of Tropical Agriculture*, 49, 2011.

[190] L. Cui, L. Li, A. Zhang, G. Pan, D. Bao, and A. Chang. Biochar amendment greatly reduces rice cd uptake in a contaminated paddy soil: a two-year field experiment. *BioResources*, 6(3):2605–2618, 2011.

[191] L. Cui, G. Pan, L. Li, J. Yan, A. Zhang, R. Bian, and A. Chang. The reduction of wheat cd uptake in contaminated soil via biochar amendment: A two-year field experiment. *BioResources*, 7(4), 2012.

[192] T. Islami, B. Guritno, N. Basuki, A. Suryanto, et al. Biochar for sustaining productivity of cassava based cropping systems in the degraded lands of east java, indonesia. *Journal of Tropical Agriculture*, 49:31–39, 2011.

[193] RD Lentz and JA Ippolito. Biochar and manure affect calcareous soil and corn silage nutrient concentrations and uptake. *Journal of Environmental Quality*, 41 (4):1033–1043, 2012.

[194] TJ Kinney, CA Masiello, B. Dugan, WC Hockaday, MR Dean, K. Zygourakis, and RT Barnes. Hydrologic properties of biochars produced at different temperatures. *Biomass and Bioenergy*, 2012.

[195] S.O. Tagoe, T. Horiuchi, and T. Matsui. Preliminary evaluation of the effects of carbonized chicken manure, refuse derived fuel and K fertilizer application on the growth, nodulation, yield, N and P contents of soybean and cowpea in the greenhouse. *African Journal of Agricultural Research*, 3(11):759–774, 2008.

[196] Y. Chen, Y. Shinogi, and M. Taira. Influence of biochar use on sugarcane growth, soil parameters, and groundwater quality. *Australian Journal of Soil Research*, 48 (7):526–530, 2010.

[197] M Hossain. Agronomic properties of wastewater sludge biochar and bioavailability of metals in production of cherry tomato (Lycopersicon esculentum). *Chemosphere*, 78(9):1167–1171, 2010. ISSN 0045-6535.

[198] KC Uzoma, M. Inoue, H. Andry, H. Fujimaki, A. Zahoor, and E. Nishihara. Effect of cow manure biochar on maize productivity under sandy soil condition. *Soil Use and Management*, 2011.

[199] E. I. Wisnubroto, M. Hedley, K. Hina, and M. Camps-Arbestain. The use of biochar from biosolids on waitarere sandy soils: effect on the growth of rye grass. Conference poster, 2010.

[200] T. Suppadit, V. Kitikoon, A. Phubphol, P. Neumnoi, et al. Effect of quail litter biochar on productivity of four new physic nut varieties planted in cadmium-contaminated soil. *Chilean Journal of Agricultural Research*, 72:1, 2012.

[201] J.H. Park, G.K. Choppala, N.S. Bolan, J.W. Chung, and T. Chuasavathi. Biochar reduces the bioavailability and phytotoxicity of heavy metals. *Plant and soil*, 348 (1):439–451, 2011.

[202] K.Y. Chan, ZH Xu, J. Lehmann, and S. Joseph. Biochar–nutrient properties and their enhancement. In *Biochar for Environmental Management: Science and Technology*, pages 67–84. 2009.

[203] C.I. Kammann, S. Linsel, J.W. Gößling, and H.W. Koyro. Influence of biochar on drought tolerance of Chenopodium quinoa Willd and on soil–plant relations. *Plant and Soil*, pages 1–16, 2011.

[204] B.L. Liang, J. Solomon, D. Kinyangi, J. Grossman, J. O'Neill, B. Skjemstad, JO Thies, J. Luizão, FJ Petersen, J. Neves, et al. Black carbon increases cation exchange capacity in soils. *Soil Science Society of America Journal*, 70(5):1719, 2006.

[205] S. Wood. mgcv: Mixed gam computation vehicle with gcv/aic/reml smoothness estimation. Technical report, 2012.

[206] Paul D Allison and Richard P Waterman. Fixed–effects negative binomial regression models. *Sociological methodology*, 32(1):247–265, 2002.

Appendix A

Heterogeneous global crop yield response to biochar: a meta-regression analysis: *Supplementary Information*

A.1 Summary Statistics

Summary statistics for soil, biochar, and management variables in our dataset are given in figures A.1, A.2, and A.3.

A.2 Statistical Methods

A.2.1 Missing data methods

Many of the studies in our dataset do not report values of many of the variables that we use (figure A.4). Standard model-fitting techniques cannot function in the presence of missing data, so statistical software packages normally omit cases of missing data. This greatly reduces statistical power where missing data are prevalent, or, in the case where commonly missing variables are dropped from the analysis, can lead to omitted variables bias. However, well-developed methods exist for drawing inference about responses to variables with partially-missing values.

We use multiple imputation (MI), a technique pioneered by Rubin [51, 63]. In MI, statistical models are specified to predict the missing values of variables, and random draws are taken from the predictive distribution of these fitted models to "fill in" missing data.

FIGURE A.1: Summary statistics for soil variables. Diagonals give kernel density plots of observed (imputed) single variables in red (blue), with their intersection in purple. Response ratio, the logarithm of the variance of response ratio are represent mean values of those variables by each unique soil. Off-diagonals give bivariate scatterplots of two variables, and lowess regression lines, with red dots indicating observed variables and semi-transparent green dots representing imputed values. Solid black lowess fits give correlations between observed values of variables, while dotted blue lines give correlations between imputed variables. RR = response ratio, SOC = soil organic carbon (g kg^{-1}), CEC = cation exchange capacity (cmol$_c$$^{-1}$), sand and clay given in percentage by mass.

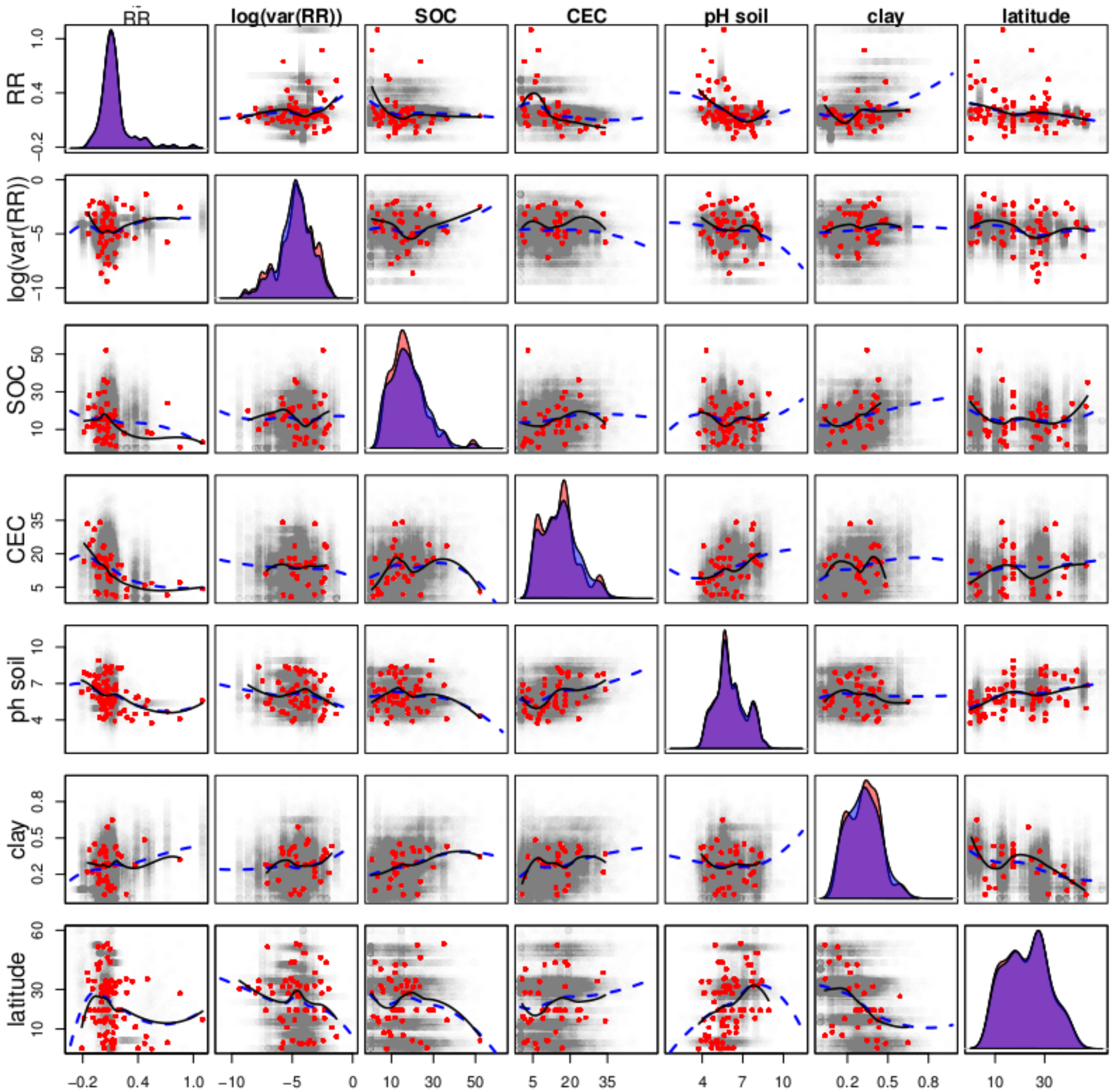


FIGURE A.2: Summary statistics for biochar variables. Diagonals give kernel density plots of observed (imputed) single variables in red (blue), with their intersection in purple. Response ratio, the logarithm of the variance of response ratio are represent mean values of those variables by each unique soil. Off-diagonals give bivariate scatterplots of two variables, and lowess regression lines, with red dots indicating observed variables and semi-transparent green dots representing imputed values. Solid black lowess fits give correlations between observed values of variables, while dotted blue lines give correlations between imputed variables. RR = response ratio, HTT = highest treatment temperature (maximum temperature attained during pyrolysis), %C = percent carbon by mass.

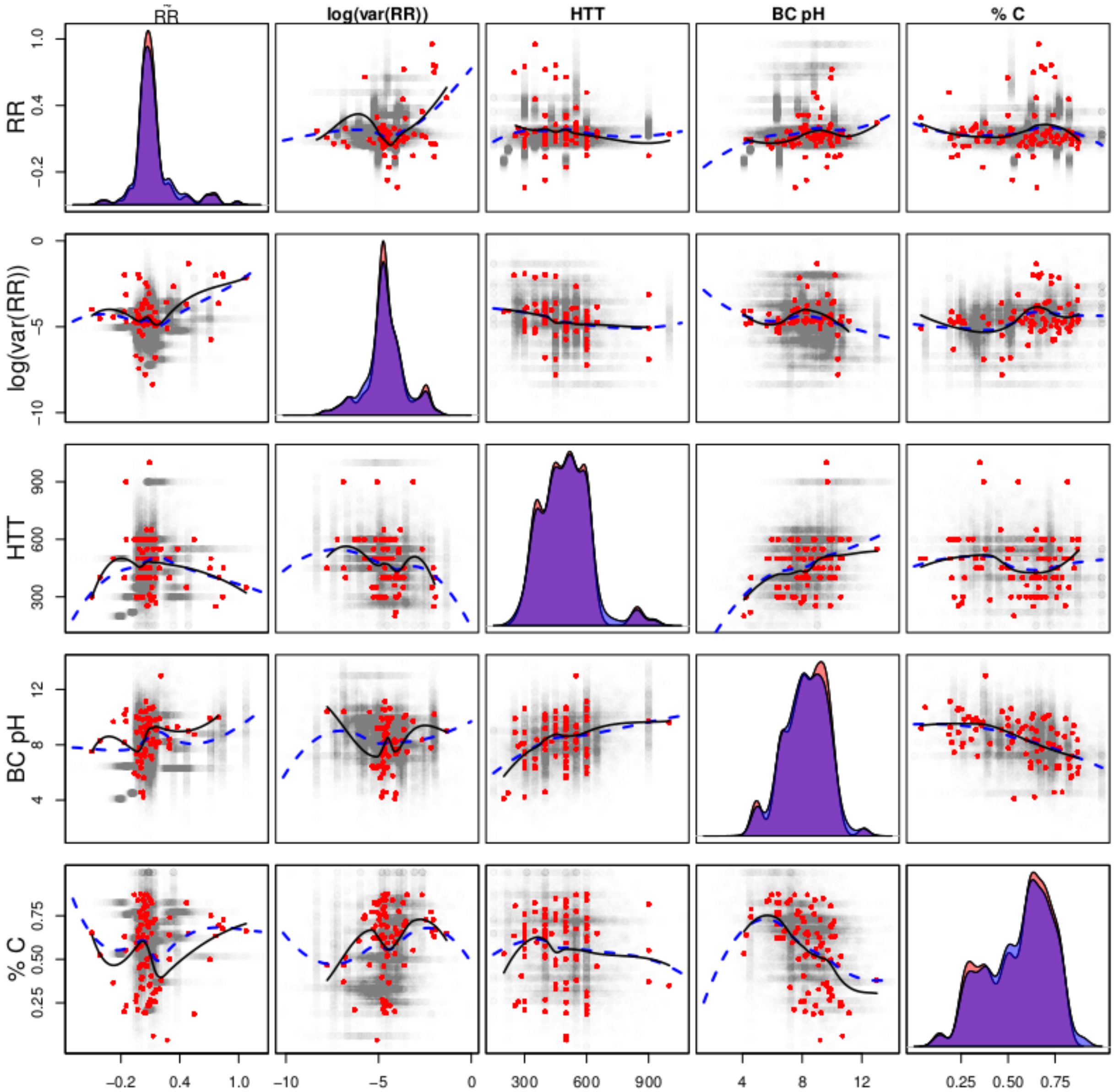


FIGURE A.3: Summary statistics for management variables. Diagonals give kernel density plots of observed (imputed) single variables in red (blue), with their intersection in purple. Response ratio, the logarithm of the variance of response ratio are represent mean values of those variables by each unique soil. Off-diagonals give bivariate scatterplots of two variables, and lowess regression lines, with red dots indicating observed variables and semi-transparent green dots representing imputed values. Solid black lowess fits give correlations between observed values of variables, while dotted blue lines give correlations between imputed variables. RR = response ratio, BCAR = biochar application rate (tons ha^{-1}), INAR = nitrogen application rate (kg ha^{-1}).

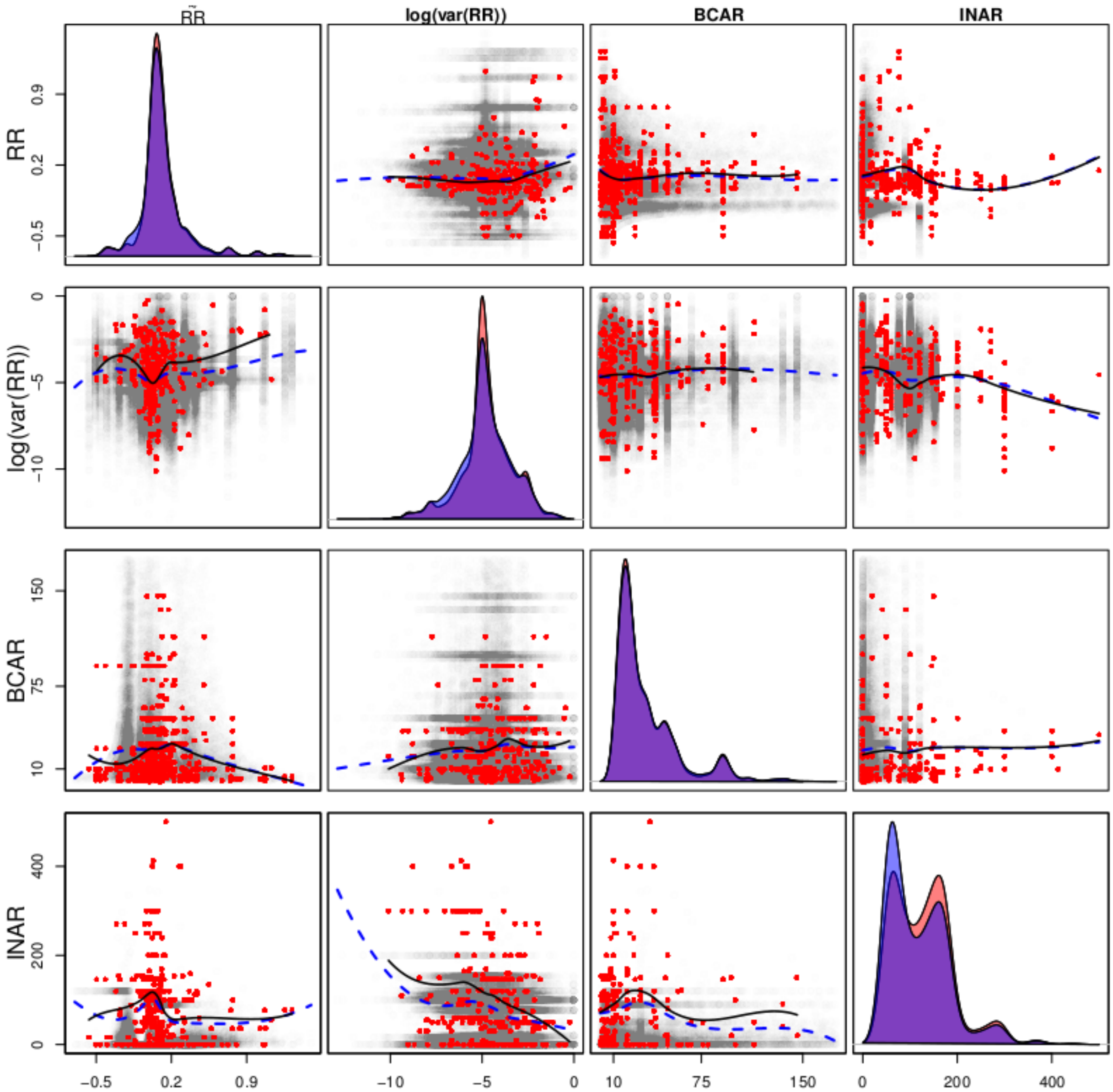
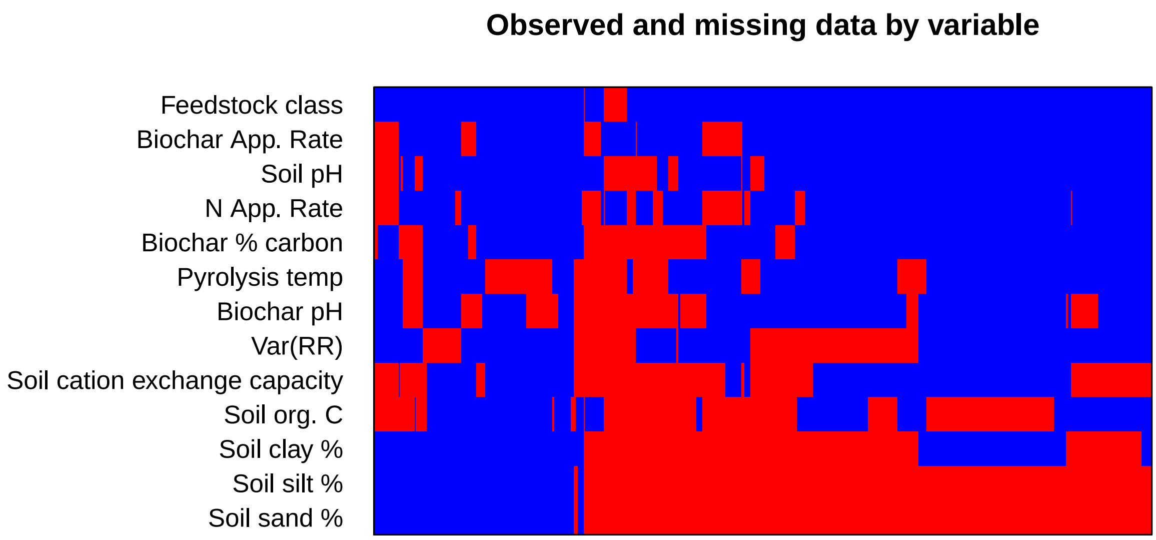


Figure A.4: Patterns of reported and missing data in our dataset of 781 response ratios from 78 studies, including observations of crop yield and biomass production. Columns indicate individual observations, blue indicates observed data, red indicates missing data.



These completed datasets are then analyzed using methods developed for datasets with no missing values. This process is repeated several times, yielding several fitted models. The results from the fitted models are then combined. Specifically, parameter estimates $\hat{\beta}_m$ from M imputations are simply averaged: $\hat{\beta} = \frac{1}{m}\sum_{m=1}^{M}\hat{\beta}_m$. The variance-covariance matrix $\hat{V}_\beta$ of the estimated parameters is calculated by first averaging variance-covariance matrices, and then adding a correction to account for variation between imputation models:

$$\hat{V}_{\hat{\beta}} = W + \left(1 + \frac{1}{M}\right) B \tag{A.1}$$

where $W = \frac{1}{M}\sum_{m=1}^{M}\widehat{VCV}_m$, $\widehat{VCV}$ is the estimated variance-covariance matrix of the estimated parameters, and $B = \frac{1}{M-1}\sum_{m=1}^{M}\left(\hat{\beta}_m - \hat{\bar{\beta}}\right)\left(\hat{\beta}_m - \hat{\bar{\beta}}\right)^T$. This procedure inflates the standard errors on coefficients about which the imputation model is relatively less certain, either due to a lot of missing data, or due to a poorly-informative imputation model.

Data can be missing randomly or non-randomly. When data is missing-not-at-random, special care must be taken to model the probability of data being missing, and to control for this factor when predicting missing data. We proceed under the assumption that data is missing at random. This assumption is justified because of the nature of our data – studies measured or did not measure various variables based on their salience to the particular hypotheses of the individual study. In the social sciences (for which multiple imputation was first developed and applied [51]), missingness of data was commonly due to non-response in surveys, and particular socioeconomic groups with different characteristics were often more or less likely to respond. Taking our data to be missing-not-at-random would require postulating that characteristics of studies that affect that affect their outcomes are correlated with the probability that data is missing.

For imputation modeling (i.e.: modeling variables and predicting their missing values), we follow Gelman & Hill [65] in using iterative regression imputation (also known as multiple imputation by chained equations or MICE [64]). We first divide our variables into four classes – soil variables, biochar variables, management variables, and outcome variables. Within each class, we shink the dataset to reflect unique observations – for example, unique chars or soils rather than repeated observations of the same soil. We assume that biochar variables are not predictors of the characteristics of the soils to which they were applied, vice versa, and that management decisions (amount of biochar added, and amount of nitrogen fertilizer added) were likewise not predictors of the types of soil to which they were applied or the characteristics of the biochar added. We believe that this assumption is reasonable because of the nature of our dataset as observations from randomized controlled trials. Were our dataset comprised of actual use of biochar by farmers, this assumption would be far more dubious, as amount of biochar or fertilizer added may well be a predictor of soil type, for example. Outcome variables include $\tilde{RR}_{ic,BC=3}$ – the estimated biochar application rate at 3 Mg ha^{-1}, and the variance of the response ratio. These variables are used as predictors in each of the three other imputation classes that we define. $\tilde{RR}_{ic,BC=3}$ is treated as a latent variable, inasmuch as it is observed directly for only a small subset of our observations that actually added biochar at that level. Its estimation, as well as the estimation of other missing variables, is described below.

Random imputations

We being by filling in missing values of variables using random draws (with replacement) from observed values of those variables. Mostly, these variables form a continuous distribution with some support specific to the variable. Biochar feedstock, as a factor variable, was drawn from it's three-point observed distribution. Nitrogen fertilizer application rate is naturally left-censored at zero, so we first imputed it as (0/>0) binomial random variable, and then imputed predicted non-zero cases as with other continuous variables.

$\tilde{RR}_{ic,BC=c}$ was imputed in the first stage using an explicit model. Our preferred specification was

$$RR = \beta(log(\text{BC. App. Rate} + 1) \times expid) + \epsilon \tag{A.2}$$

where *expid* is a factor variable for common-control observations. This specification simply fits a log-linear slope through each set of common-control response ratios, running through the origin. This is motivated by the nature of the dependent variable as a response ratio of a biochar treatment to a non-biochar control – its expectation must be zero when no biochar is added. In addition to this one, we explore several alternative specifications (section 3, below).

Soil imputations
We use the random imputations as a starting point for explicit models predicting missing values of variables. We begin with soil variables, by first creating a temporary dataset wherein each unique soil is represented only once, thereby shrinking the size of the dataset and avoiding a counter-intuitive situation whereby identical soils might be estimated to have different characteristics. Models are specified as follows:

$$Y^{miss} = X'^{imputed}\beta + \gamma_1\overline{\tilde{RR}}_{BC=3} + \gamma_2 log(\overline{varRR}^{imputed}) + \epsilon \tag{A.3}$$

where *miss* indicated a partially-missing vector and *imputed* indicates a vector comprised of observed data where available, and imputed data where missing. The unique-soil-wise means of the response ratio and its variance were also controlled for, as failure to include outcome data in imputation models is known to bias their output.

We cycle through our available soil variables, alternately specifying Y as soil pH, cation exchange capacity, soil organic carbon, clay content, sand content, and the absolute value of the study's lattitude. Variables specified as dependent variables are removed from X at each step.

After fitting each model, we take a draw from its predictive distribution for each missing datapoint, which we use to update imputed values of partially-missing variables. After cycling through each of the 6 soil variables, we merge the updated soil imputations sub-dataset back into the main dataset.

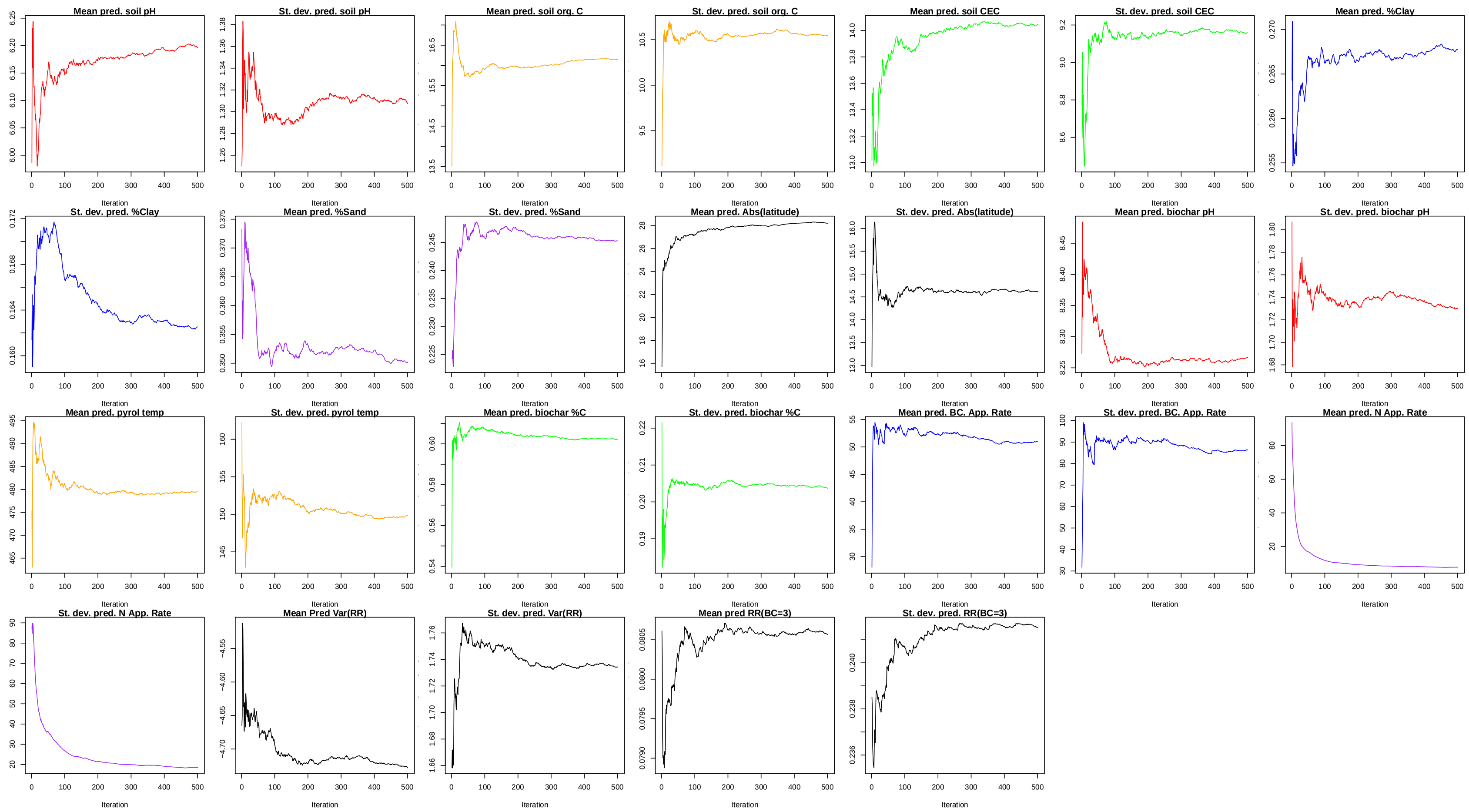
Biochar imputations are conducted in much the same way. Variables cycled through include biochar pH, biochar percentage carbon, pyrolysis temperature, and feedstock class (wood, nonwood, or manure). For feedstock class, we use a multinomial logit model, rather than an OLS regression.

Imputations for management variables are conducted somewhat differently. For these (biochar application rate and nitrogen application rate), we use the full dataset, and account for their left-censored nature. For biochar application rate, we model its logarithm as a function of our soil and biochar variables, outcome variables, and interactions between our outcome variables and crop type, as well as field/greenhouse trial as factor variables. For nitrogen application rate, we first estimate a logistic regression of whether or not N was added, and where our posterior draw predicts that N was added, estimate the logarithm of nitrogen application rate as a linear function of the same variables as those predicting biochar application rate. The difference in approach between the two variables is motivated by the large number of unfertilized observations in our dataset, while all observations added biochar in some non-zero quantity.

Imputations for outcome variables: the variance of the response ratio is predicted similarly to the biochar application rate, using all available explanatory variables and a logarithmic representation of the (highly skewed) dependent variable. $\tilde{RR}_{ic,BC=c}$ was then re-estimated using model A.2, with updated imputations for biochar application rate.

Iteration and convergence: the steps described in this section were completed sequentially, updating imputations at each step. The entire algorithm was repeated many times – further updating imputations – until approximate convergence in distribution was reached (i.e.: until the mean and variance of each of the several hundred imputations, taken together, approached some value). An example of a convergence diagnostic plot is given in figure A.5. Approximate convergence is achieved after 200 cycles (or, at minimum, oscillations/changes diminish to exceedingly small practical magnitude). This is the number of times that we run the algorithm for the production of the multiple imputed datasets to which we fit our model of interest.

FIGURE A.5: A convergence plot from a single imputation via chained equations. X-axes display iterations of the algorithm, and colored lines indicate the evolution of the means and variances of all sets of imputed values.



Production of multiple datasets is accomplished simply by running the above-described algorithm 200 times (in which 200 imputation cycles are conducted, each), to generate $M = 200$ imputed datasets. We use such a large number of imputations to engender numerical stability and reproducibility of our results[64, 173, 174].

A.3 Different specifications for estimating $\tilde{RR}_{is,BC=c}$

We expect that biochar application rate influences plant outcomes, and we further expect that this influence would be heterogeneous across biochars, soils, and perhaps other factors. However, we are unable to control for the biochar application rate as a simple additive term in our main statistical model because of the nature of our dependent variable as a response ratio. Including the amount of biochar as a simple additive term would lead to the non-zero estimates of response rate at zero biochar application. We therefore construct a dependent variable representing an estimated response ratio for a common level of biochar application rate; response ratio at for 3 Mg ha^{-1}. Doing so involves estimating heterogeneous slopes mapping plant response to biochar application rate for different studies.

We investigate 6 potential specifications for estimating these slopes. Each specification constrains each slope to run through the origin, reflecting that response ratio equals 0 when biochar application rate equals zero, by definition. The first two specify linear (model A.4) or logarithmic (model A.5) slopes between response ratio and biochar application rate for each set of common-control observations:

$$RR = \beta(\text{BC. App. Rate} \times expid) + \epsilon \tag{A.4}$$

$$RR = \beta(log(\text{BC. App. Rate} + 1) \times expid) + \epsilon \tag{A.5}$$

where *expid* represents a set of common-control observations, which will be a set with one or more member. In sets with one member, these equations draw a linear slope between the observation and the origin, or between the average of the members of the set and the origin in cases where there is more than one observation sharing a common control.

In addition to these linear fixed-effects slopes, we also estimate analogous random coefficients models specifying linear and logarithmic slopes, with and without interactions between soil clay content and soil cation exchange capacity. In specifications with interactions, we allow the size of the slope to be mediated by soil clay content and soil cation exchange capacity, as it is possible that clay may form biochar-mineral complexes [77], in ways that may plausibly be mediated by the surface chemistry of constituent clays.

$$RR = \left(\beta^{all} + \beta^{exp}\right) \text{BC. App. Rate} + \epsilon \tag{A.6}$$

$$RR = \left(\beta^{all} + \beta^{exp}\right) log(\text{BC. App. Rate} + 1) + \epsilon \tag{A.7}$$

$$\begin{aligned} RR = & \left(\beta^{all} + \beta^{exp}\right) \text{BC. App. Rate} \\ & + \gamma_1 \left(\text{BC. App. Rate} \times Clay\%\right) + \gamma_2 \left(\text{BC. App. Rate} \times CEC\right) + \epsilon \end{aligned} \tag{A.8}$$

$$\begin{aligned} RR = & \left(\beta^{all} + \beta^{exp}\right) log(\text{BC. App. Rate} + 1) \\ & + \gamma_1 log((\text{BC. App. Rate} + 1) \times Clay\%) + \gamma_2 log((\text{BC. App. Rate} + 1) \times CEC) + \epsilon \end{aligned} \tag{A.9}$$

Equations A.6, A.7, A.8, & A.9 treat the slopes as a sample from a normally-distributed population of slopes, potentially mediated by clay content and soil cation exchange capacity.

Estimates of $\tilde{RR}_{BC=3}$ (transformed into relative increase) from equations A.4,A.5, A.6, A.7, A.8, & A.9 are given in figure A.6, and fits to the data from each of these models are given in figures A.7, A.8, A.9, A.10, A.11 & A.12. Figure A.8 gives the best visual fit to data, and the most favorable AIC of the specificaitons tested. In particular, random coefficients models estimate slopes that tend towards the mean slope for the data. Of the two fixed slope models, we use the logarithmic one to drive calculate $\tilde{RR}_{BC=3}$ to serve as the dependent variable for model A.10.

Figure A.6: Estimates of relative increase from equations A.4,A.5, A.6, A.7, A.8, & A.9. Untransformed relative increase, as reported in primary studies, given in black.

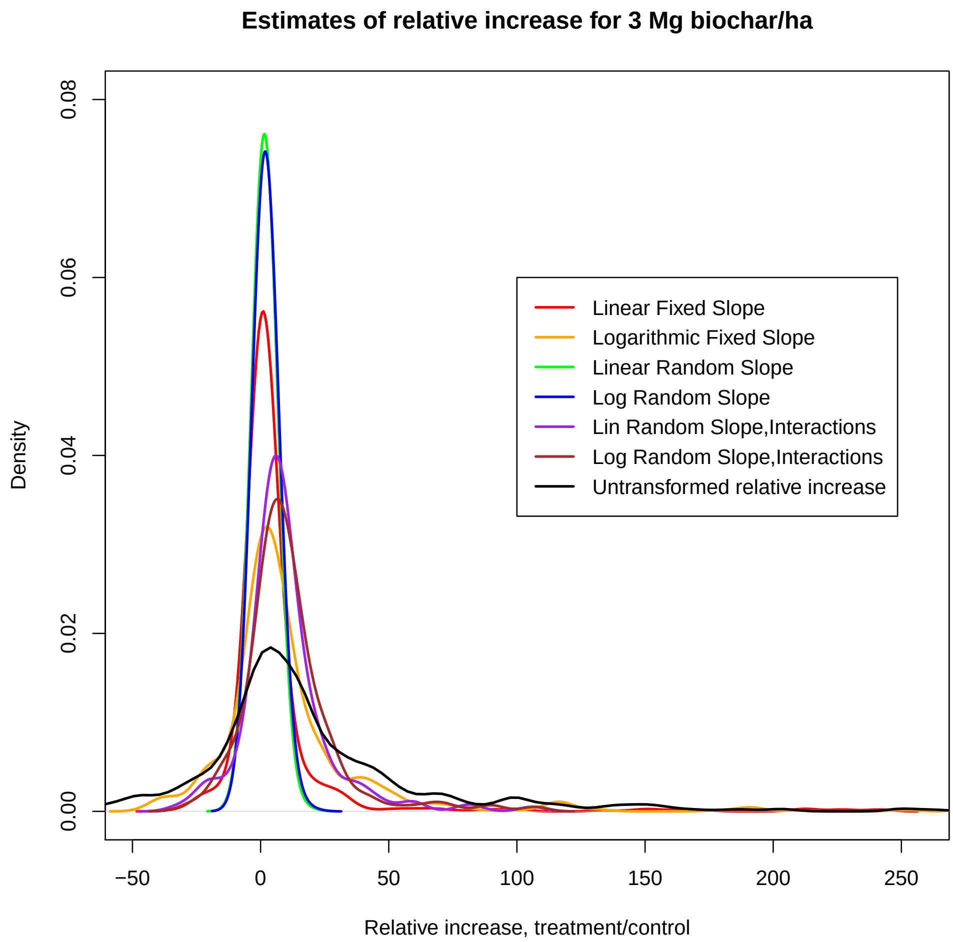


Figure A.7: Slope estimates from equation A.4. Dots indicate observed values, green asterisk indicate predicted values, and blue lines indicate their divergence. Left and right figures show the same data; the right figure is zoomed into the biochar application rates < 20 Mg ha^{-1}

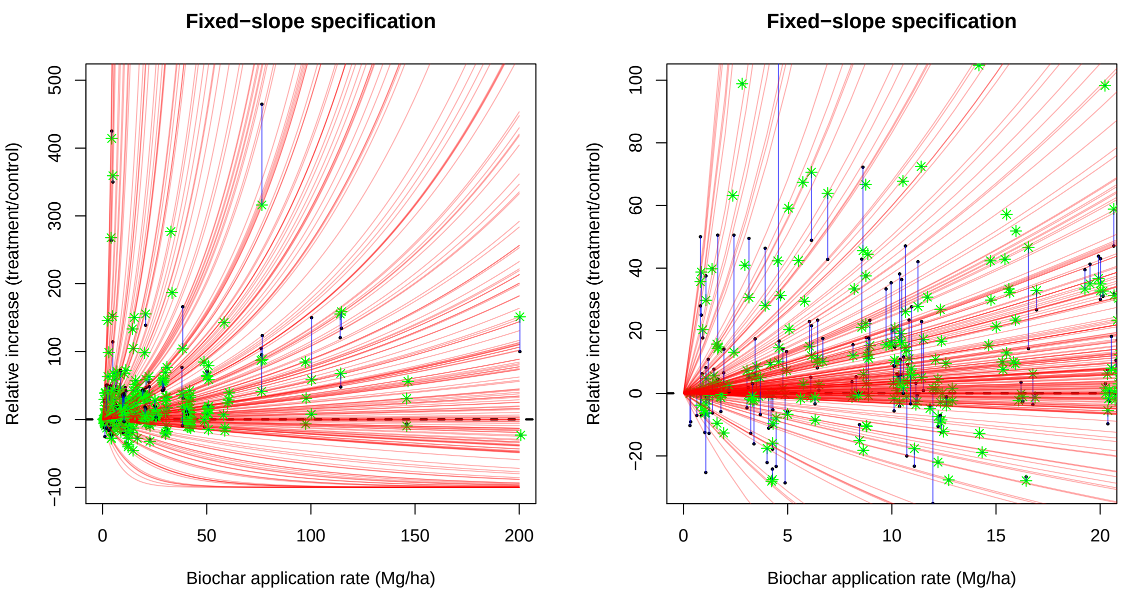


Figure A.8: Slope estimates from equation A.5. Dots indicate observed values, green asterisk indicate predicted values, and blue lines indicate their divergence. Left and right figures show the same data; the right figure is zoomed into the biochar application rates < 20 Mg ha^{-1}

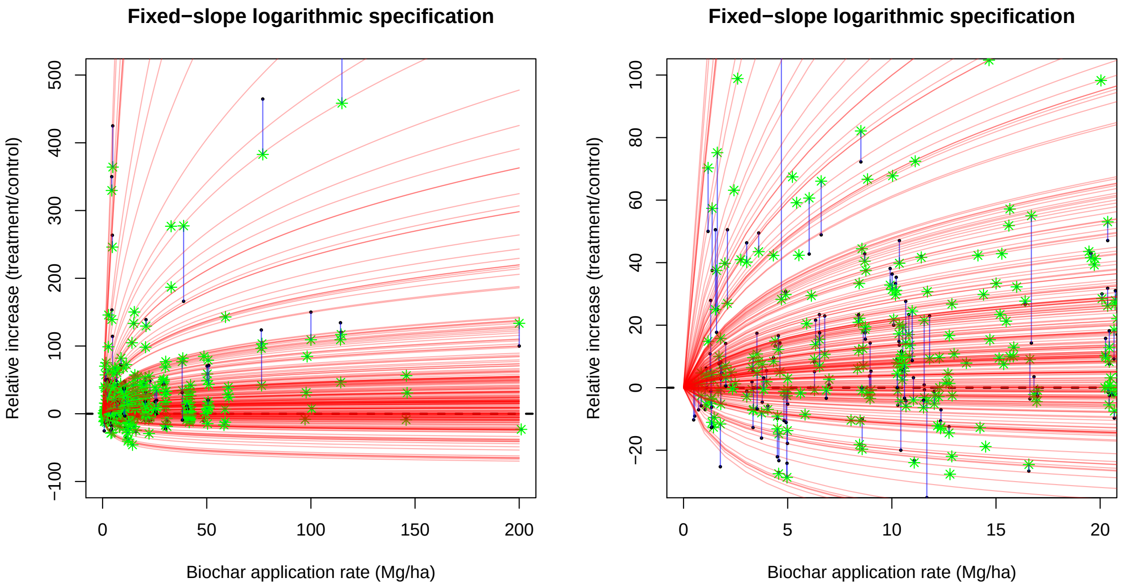


FIGURE A.9: Slope estimates from equation A.6. Dots indicate observed values, green asterisk indicate predicted values, and blue lines indicate their divergence. Left and right figures show the same data; the right figure is zoomed into the biochar application rates < 20 Mg ha^{-1}

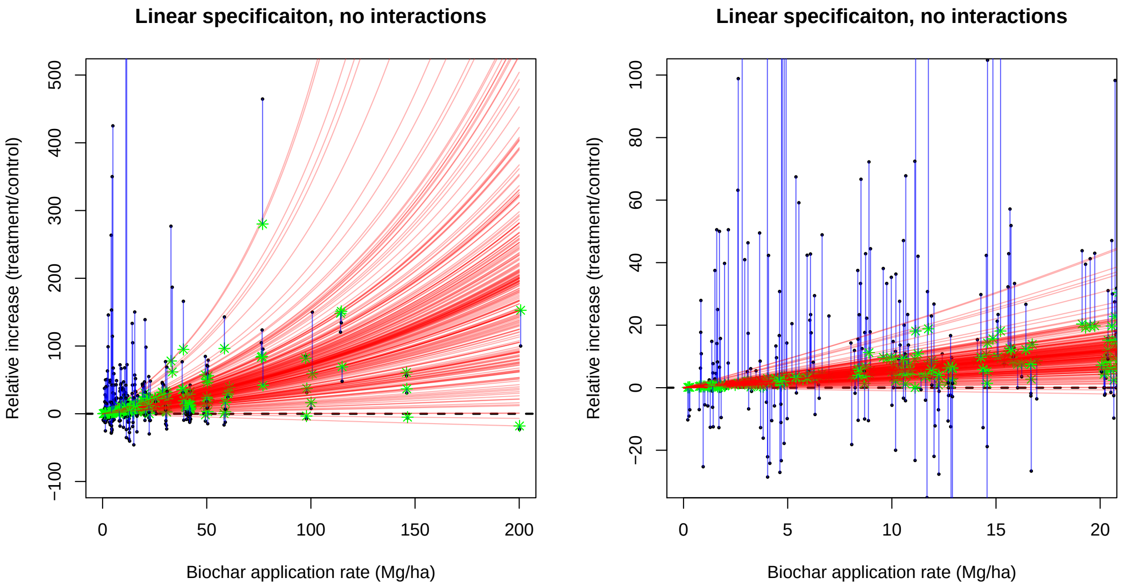


FIGURE A.10: Slope estimates from equation A.7. Dots indicate observed values, green asterisk indicate predicted values, and blue lines indicate their divergence. Left and right figures show the same data; the right figure is zoomed into the biochar application rates < 20 Mg ha^{-1}

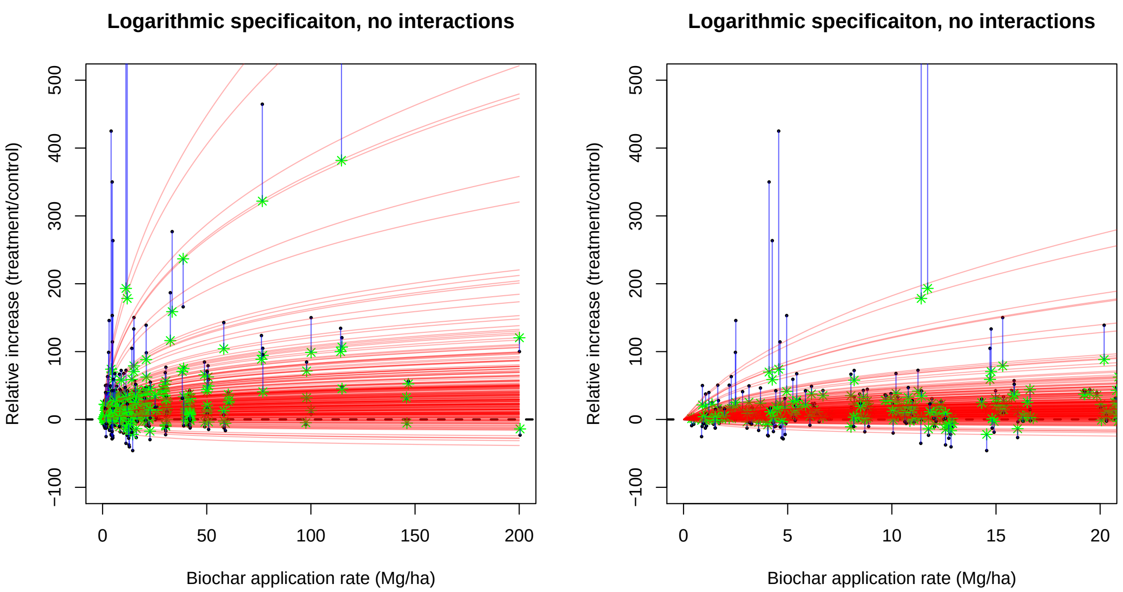


FIGURE A.11: Slope estimates from equation A.8. Dots indicate observed values, green asterisk indicate predicted values, and blue lines indicate their divergence. Left and right figures show the same data; the right figure is zoomed into the biochar application rates < 20 Mg ha^{-1}

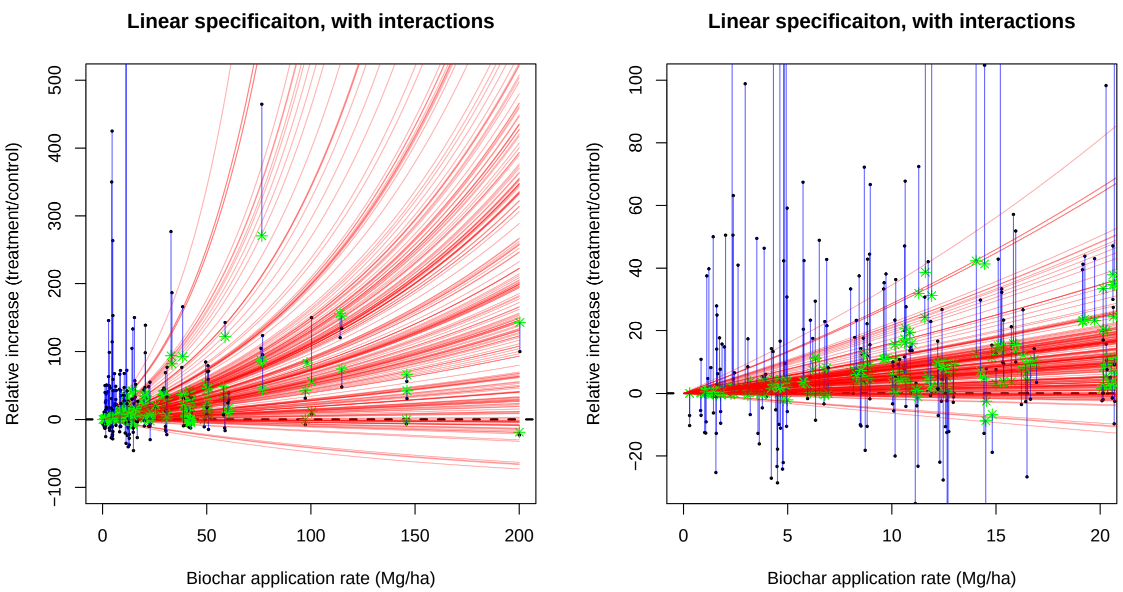
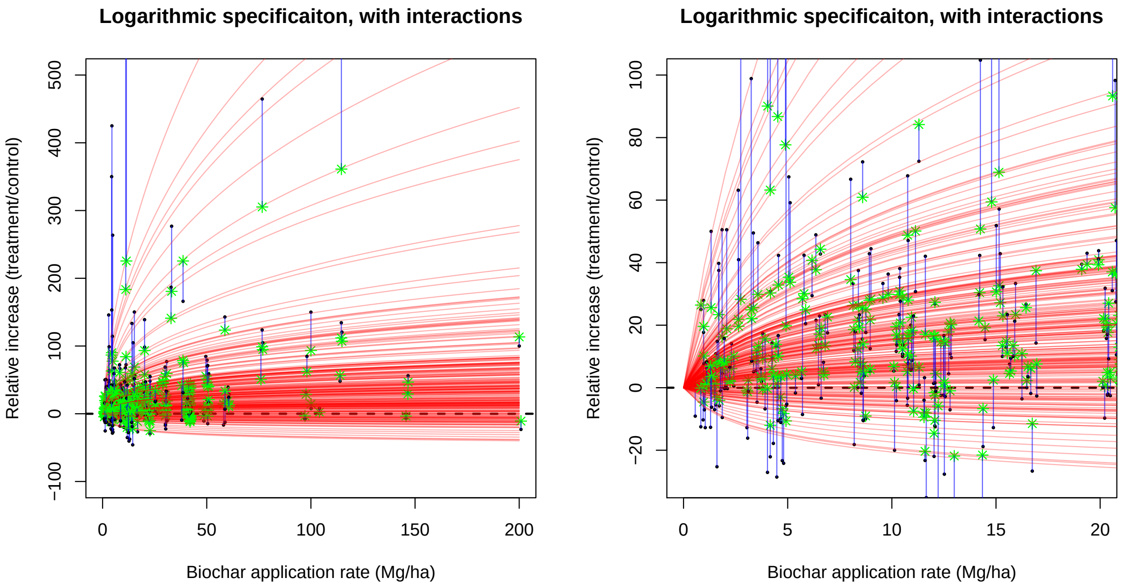


Figure A.12: Slope estimates from equation A.9. Dots indicate observed values, green asterisk indicate predicted values, and blue lines indicate their divergence. Left and right figures show the same data; the right figure is zoomed into the biochar application rates < 20 Mg ha^{-1}



A.4 Model goodness-of-fit by study

Figure A.13 gives a diagnostic of how well individual studies were fit by model A.10. The largest median divergence was less than 0.1 response ratio units, which is approximately 11 percentage points of relative increase. Residuals for several studies were highly variable, indicating poor ability to predict their missing values in imputation models, and/or large heterogeneity in within-study covariates. Note that a relatively larger portion variance explained by the model was explained by unique soil-biochar random effects than by our explanatory fixed effects, which is to say that the between-context variation unattributable to the soil, biochar and management factors that we include in our model was more important in explaining response ratio than were those factors.

FIGURE A.13: Goodness of fit measured by the distribution of the residuals from each fo the 200 fitted models for each datapoint, by study. Studies with the bulk of the distribution of their residuals below the dotted red line were negative outliers, and studies with the bulk of their distribution above were positive outliers. Distributions are formed from residuals from each of the 200 imputations used to fit our model. The x axis lists the last name of the lead author of the study.

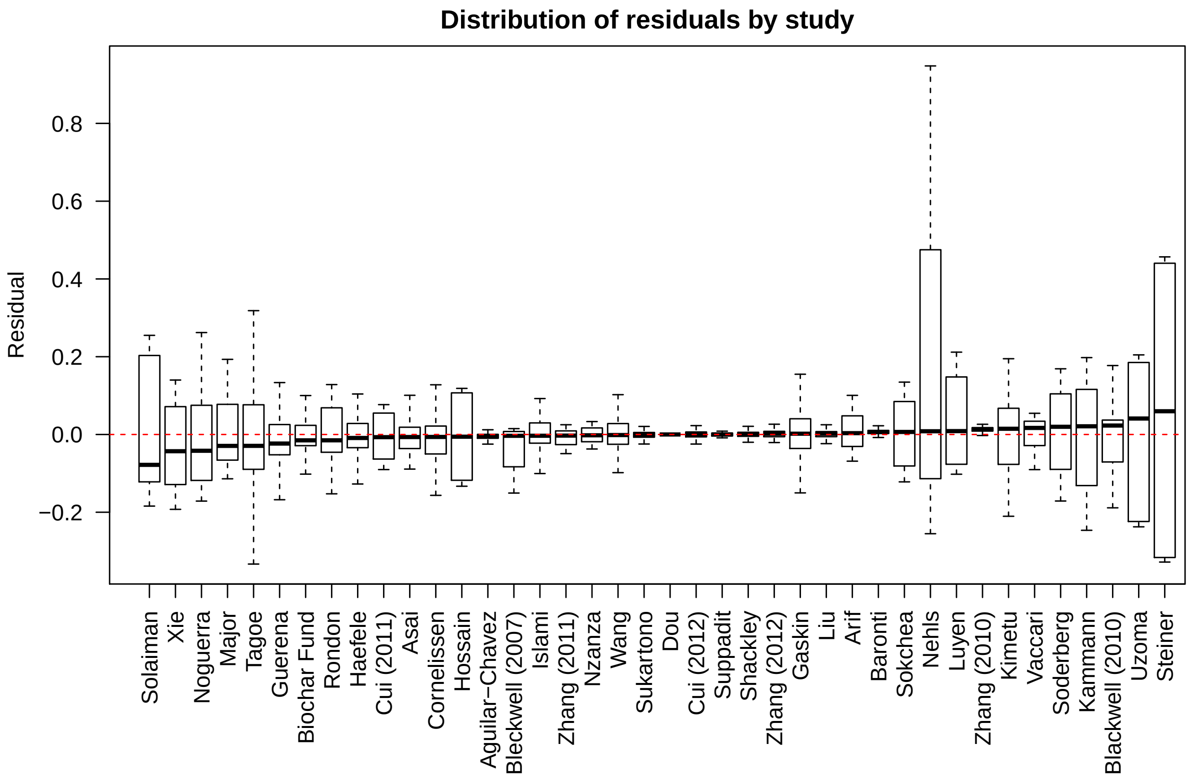
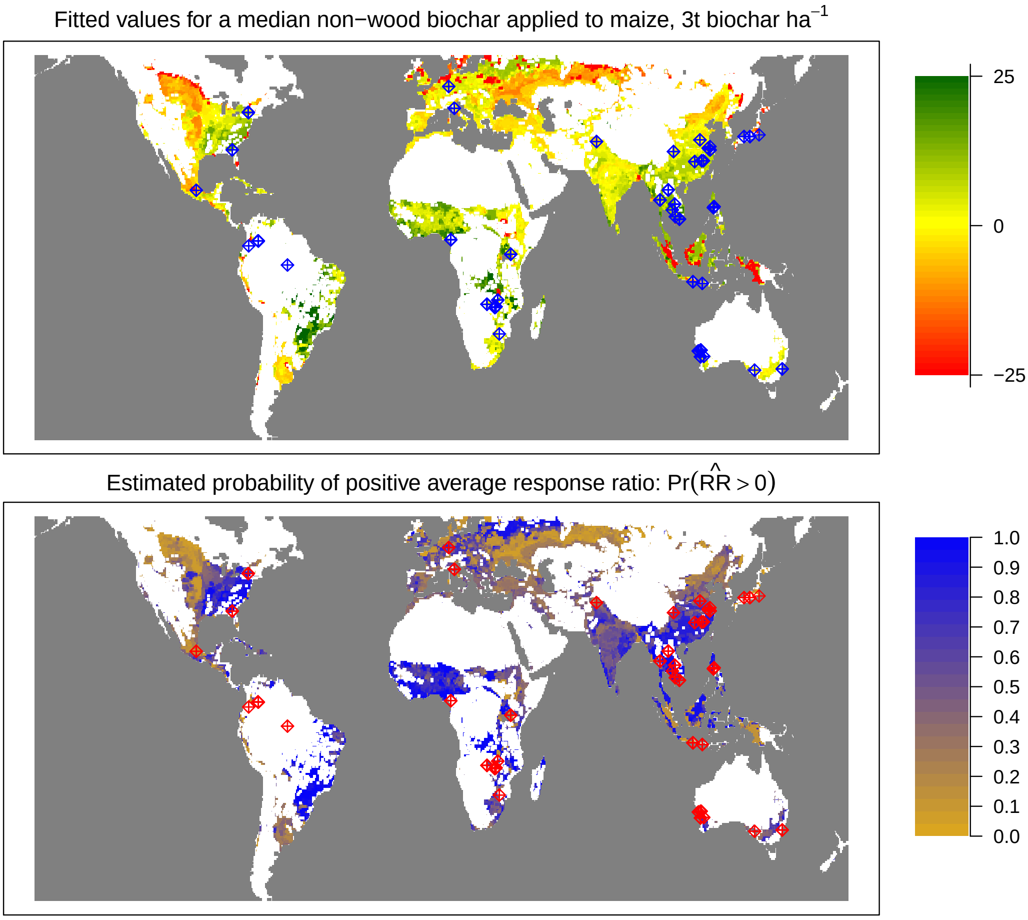


FIGURE A.14: Model predictions for a median non-wood biochar applied to maize, projected onto spatially-averaged soil property dataset from Batjes (2005)[55]. Upper maps give point estimates of percentage relative increase over control for a .5x.5 degree gridcell, and lower maps give the statistical confidence that the estimated fitted value is greater than zero. Areas with < 5% of land in agriculture are masked using agricultural area extent data from Ramankutty [67].



A.5 Standard errors of fitted values for projection maps

Our fitted model, projected onto Batjes' [55] global database of soil properties for representative biochars of our three feedstock classes, is given below in figure A.14.

A.6 Heterogeneity by crop type

By pooling response ratios derived from different crops, we are implicitly making the assumption that crop yields are moderated identically by our model terms across all crop types. For example: crop yield response to biochar in increasingly acidic soil is identical between rice and wheat. There is no *a priori* reason to believe that this is true, but we do so because our dataset is too small to treat each crop separately.

It is important to note that our main model controls for heterogeneity in average response, both through random effects for unique soil-biochar combinations, and, where more than one type of crop was grown on a unique combination of soil and biochar, indicator variables for crop types. This section explores unmodeled heterogeneity in response to moderating variables, rather than modeled heterogeneity in overall means.

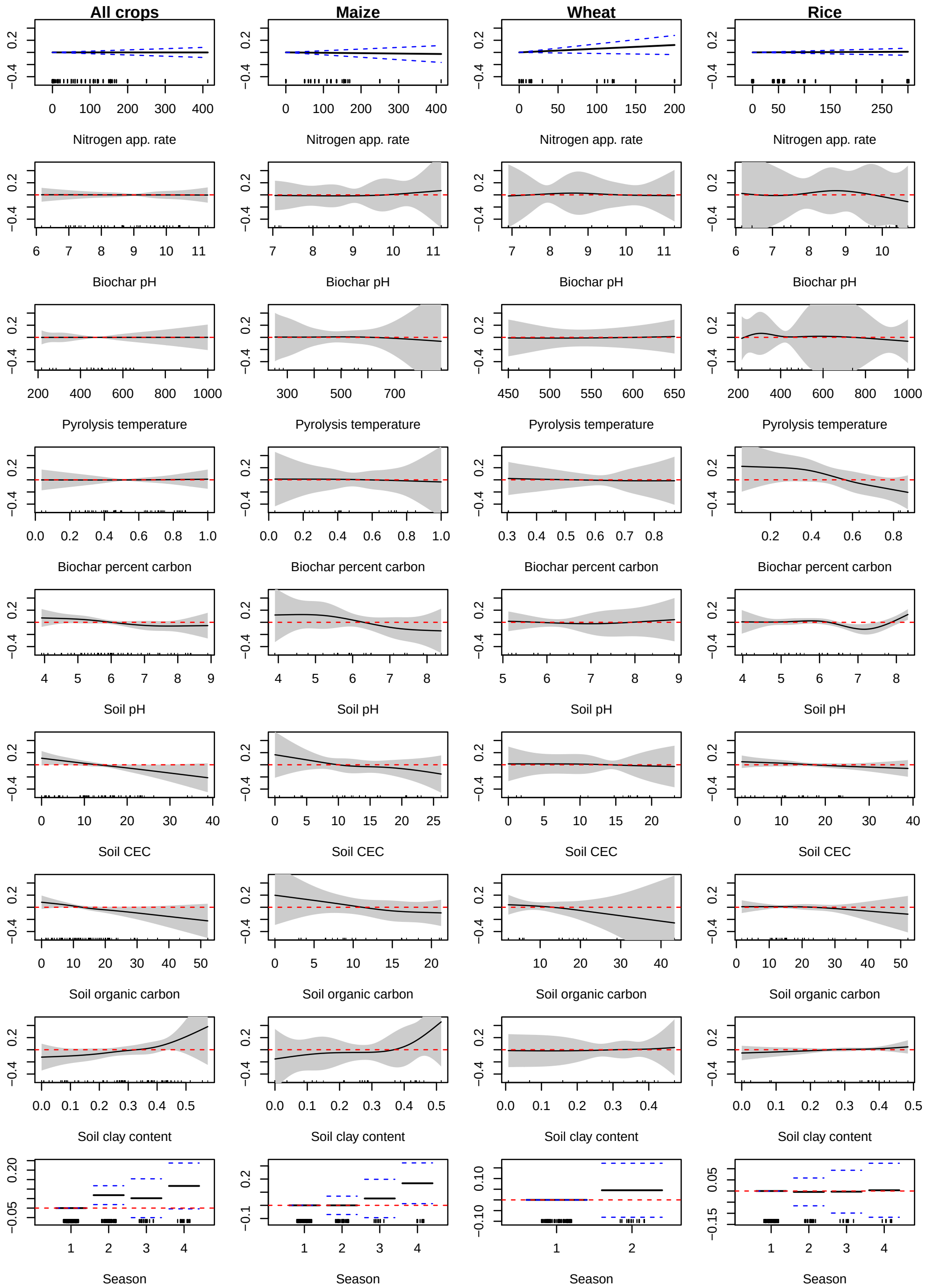
We explore this heterogeneity by fitting our model to subsets of our data (with fewer knots to accomodate fewer degrees of freedom). Estimates of response to the continuous variables and over time is given in figure A.15.

Only very modest heterogeneity is apparent – we observe suggestive evidence that rice appears to benefit from biochars with lower carbon contents, and that the response of maize to biochar is mediated to a greater degree by soil clay and soil organic carbon content than it is for other crops. However these effects are relatively modest, and difficult to distinguish from one another statistically. Heterogeneity between crops in plant response to different sorts of biochar remains an area for further study.

A.7 Model with interactions

In addition to our main predictive model, we estimate a richer model, specified in equation A.10 with a number of interaction terms. All interaction terms are specified as tensor products [162], denoted "$\otimes$."

Figure A.15: Estimated nonparametric smooth functions from a model with all of our explanatory variables but no interactions, for different subsets of our dataset defined by crop type.



$$
\begin{aligned}
\tilde{RR}_{is,BC=3\text{Mg ha}^{-1}} = \; & \alpha \\
& +\tau_s \\
& +\beta_1(\text{Pot trial}{=}{=}1) \\
& +\beta_{2-4}(\text{Seasons since initial application}) \\
& +\beta_{5-7}(\text{feedstock types}) \\
& +\beta_{6-17}(\text{Crop type}) \\
& +f(\text{N application rate}) \\
& +f(\text{Pyrolysis temperature}) \\
& +f(\text{Biochar \% carbon}) \\
& +f(\text{Biochar pH}) \\
& +f(\text{Soil cation exchange capacity}) \\
& +f(\text{Soil pH}) \\
& +f(\text{Soil organic carbon}) \\
& +f(\text{Soil \% clay}) \\
& +f(\text{Soil pH} \times \text{season since initial application}) \\
& +f(\text{Biochar pH} \times \text{season since initial application}) \\
& +f(\text{Biochar pH} \otimes \text{soil pH}) \\
& +f(\text{Soil pH} \otimes \text{soil organic carbon} \otimes \text{cation exchange capacity}) \\
& +f(\text{Clay \%} \otimes \text{cation exchange capacity}) \\
& +f(\text{Pyrolysis temperature} \otimes \text{\%C} \times \text{feedstock type}) \\
& +\epsilon_{is}
\end{aligned}
\tag{A.10}
$$

Selection of interaction terms based on the following hypotheses, drawn from the agronomic literature on biochar and from basic soil science:

1. **Yield will be maximized where alkaline biochars are applied to acidic soils**
   Biochar is commonly alkaline, and has been observed to increase soil pH in a number of studies in our dataset. Biochar's alkalinity is derived from the concentration of plant mineral fractions during pyrolysis, and both the loss of H and the formation of carbonates during pyrolysis [59]. Several studies in our dataset observe large yield responses on low-pH soils [48, 175–180], while Jeffery et al. [44] find higher average responses among the subset of studies performed on acidic soil. We expect the effect of biochar alkalinity on crop yields to vary with the pH of the soil to which it is added, as soils with neutral or high pH would benefit less from liming. This hypothesis motivates the interaction between soil pH and biochar pH.

2. **Plant response to biochar remains positive over time since application**
   Biochar is persistent in soil [28]. In spite of this persistence however, surfaces of biochar particles are both biologically and chemically reactive with properties that change over time [35, 181, 182]. Several studies in our dataset observed plant response to biochar over multiple harvest seasons [49, 68–71, 183–193], often finding that yields either increased relative to control, or that first yield increases were maintained. In addition, we hypothesize that any plant response to high-pH biochars will attenuate over time, as cations such as potassium are either leached or taken up by plants and exported at harvest. This hypothesis motivates the interaction

between soil and biochar pH and time, as well as the interaction between soil pH and biochar pH. We have too few available datapoints to construct different tensor product smooths for biochar pH $\otimes$ soil pH for each year in our dataset.

3. **Biochar's benefits are maximized in poor soils**, such as those depleted in soil organic carbon and low in cation exchange capacity, as it can provide a more stable source of many of the functions played by other forms of soil organic matter, including serving as refugia for soil biota, and as a source of amphoteric charge. Furthermore, soil CEC, soil pH and soil organic carbon content are fundamentally inter-related: the component of a soil's CEC that is derived from soil orgaic carbon is pH-dependent. Therefore, the influence of any one of these variables on crop yield response to biochar may well depend on the values of the others. We generally expect response to biochar to be high when any of them is low. This motivates the three-way interaction between soil pH, CEC, and soil organic carbon.

4. **Plant response to biochar depends on pyrolysis conditions and feedstock**. Biochar is produced by the heating of biomass in a sub-stoichiometric environment. Varying temperatures and varying degrees of oxygen occlusion will produce products with different characteristics, which in turn will influence agronomic utility. Under laboratory conditions where oxygen is completely occluded, higher temperatures will yield biochars with lower H/C and O/C ratios [58]. In addition, biochars produced at higher temperatures may be less hydrophobic than biochars produced at lower temperatures [194]. However, imperfect oxygen occlusion occurring in less sophisticated pyrolysis devices will lead to biochars with higher ash concentrations per unit carbon (because a greater fraction of the initial carbon will be lost as CO_2), depending on the mineral concentration of the initial feedstock. Biochars produced at lower temperatures may have higher concentrations of available macro-nutrients – a particular concern when dealing with charred animal waste, which has been the subject of numerous papers in our dataset [179, 189, 192, 195–201]. Loss of macro-nutrients with increasing pyrolysis temperature[202] will likely be particularly important with these chars, which are likely to have higher concentrations of nitrogen and phosphorous than chars made from vegetable matter such as wood or grass. This motivates the interaction between pyrolysis temperature, biochar percent carbon, and feedstock.

5. **Plant response to biochar will be higher in sandier soils, and in soils rich in weathered clays**.
   Biochars may improve water holding capacity, water drainage and infiltration[194, 203], and improve cation exchange[204]. These properties are typically needed most in sandy soils, and in soils whose clay constituents are highly weathered. We therefore expect that yield increases would be maximized in soils with low clay content, ans soils with high clay content but low CEC. This motivates the interaction between clay content and soil CEC.

A.7.1 Model Specification and Fit

Given the large number of interactions in model A.10, we specify only 4 knots per variable, as opposed to 10 knots per variable for the main model without interactions. Tensor product interactions in particular require k^n degrees of freedom (pre-penalization), where k is the number of knots and n is the number of terms in the interaction. We specify our knots based on a combination of subject-matter knowledge (i.e.: where we would expect breakpoints to lie) and with the goal of adequate coverage for the support of our dataset. The knots are specified as follows:

- Soil pH: 4.5, 5.5, 6.5, 7.5
- Soil cation exchange capacity: 5, 12, 19, 26 $\mathrm{cmol}_c\ \mathrm{kg}^{-1}$
- Soil organic carbon content: 10, 20, 30, 40 $\mathrm{g\ kg}^{-1}$
- Soil clay content: 15, 30, 45, 60 %
- Biochar pH: 7.5, 8.5, 9.5, 10.5
- Biochar percent carbon: 30, 45, 60, 80 %
- Pyrolysis temperature: 350, 450, 600, 800 °C

Model A.10 fits the data relatively well, with R^2 averaging .87. It is significantly different from a model with only random effects at F=7.1 ($p < .01$), and significantly different from our main model at F=10.5 ($p < .01$). While the model as a whole is significant and achieves a favorable AIC compared to our main model, none of the individual terms are statistically significant individually.

A.7.2 Results

Because none of the terms are individually significant, we present contour plots of the fitted values of the model for most terms, without measures of statistical confidence. These are intended to be purely suggestive, as we are unable to rule out the hypothesis that any of our terms individually are different from zero.

Figure A.16 gives estimated yield response to biochar over gradients of soil pH, biochar pH, and time. Results are qualitatively similar to univariate smooths given in the main results – response ratio is larger in more acidic soils, and there is little evidence that biochar pH plays much role in predicting plant response except in season 2, where lower-pH biochars are associated with higher yields. Figure A.17 gives estimated yield response to biochar over gradients of soil cation exchange capacity, soil organic carbon content, and soil pH. Yield response is maximized at low levels of soil organic carbon content and low cation exchange capacity. The interaction between these variables appears to be mediated by soil pH, which changes the shape of the functions relating cation exchange capacity and soil organic carbon to one another and to response ratio. Figure A.18 gives estimated yield response to biochar over gradients of soil cation exchange capacity, and soil clay content. Clay soils with low CEC are predicted to respond positively to biochar. Figure A.19 gives plant response to biochar for non-wood, manure, and wood biochars produced at different temperatures and having different carbon contents. While estimates are quite imprecise, there is slight suggestive evidence that higher pyrolysis temperature is not beneficial in most circumstances. The effect of biochar carbon content is most pronounced for manure chars, though these estimates are so imprecise that it is difficult to take even suggestive evidence from them.

Figure A.20 projects the fitted values of model A.10 onto Batjes' [55] global soil property database. Results are qualitatively similar though generally more positive than those for the main model given in figure A.14. The two models disagree in effect sign primarily in organic soils in the tropics, notably in Indonesia.

As judged by AIC, model A.10 is superior to our main model with no interactions. However, we are reluctant to present model A.10 in our main results for two reasons. First, no terms are individually statistically significant – due largely to collinearity among terms. As such, it is difficult to infer precisely which terms are key drivers of observed response ratios. Secondly, we specify only a handful of the interactions that could be reasonably postulated – our key constraint is available degrees of freedom. With infinite degrees of freedom, the tensor-product interactions might reasonably be specified as one very large interaction of 5 terms or more, perhaps including biochar pH, soil pH, soil cation exchange capacity, soil organic carbon, and soil clay content – potentially interacted further with time or feedstock. This would be desirable because the effect of changing any one of these variables could quite plausibly depend on the value of each of these other variables. Given that we lack the data for such a model, we prefer to present a more parsimonious specification with ready interpretation. Lastly, given that the no-interaction specification gives slightly lower estimates of response ratio, we emphasize it as a more conservative lower bound.

A.8 Data, R code, and original studies

Data and R code for implementing/reproducing our analysis is available at http://andrewcd.berkeley.edu/BCMA_rcode.zip. The password (required) to decrypt the archive is "biochar." Please contact the corresponding author with any requests for primary studies not readily accessible online.

All original studies that comprise our dataset which are available from peer-reviewed journals can be found at those outlets. Unpublished studies comprising the dataset (e.g.: conference presentations/posters, unpublished reports, masters theses) are archived

FIGURE A.16: Contour plot giving fitted values of model A.10 over gradients of soil pH, biochar pH, and number of seasons since initial application.

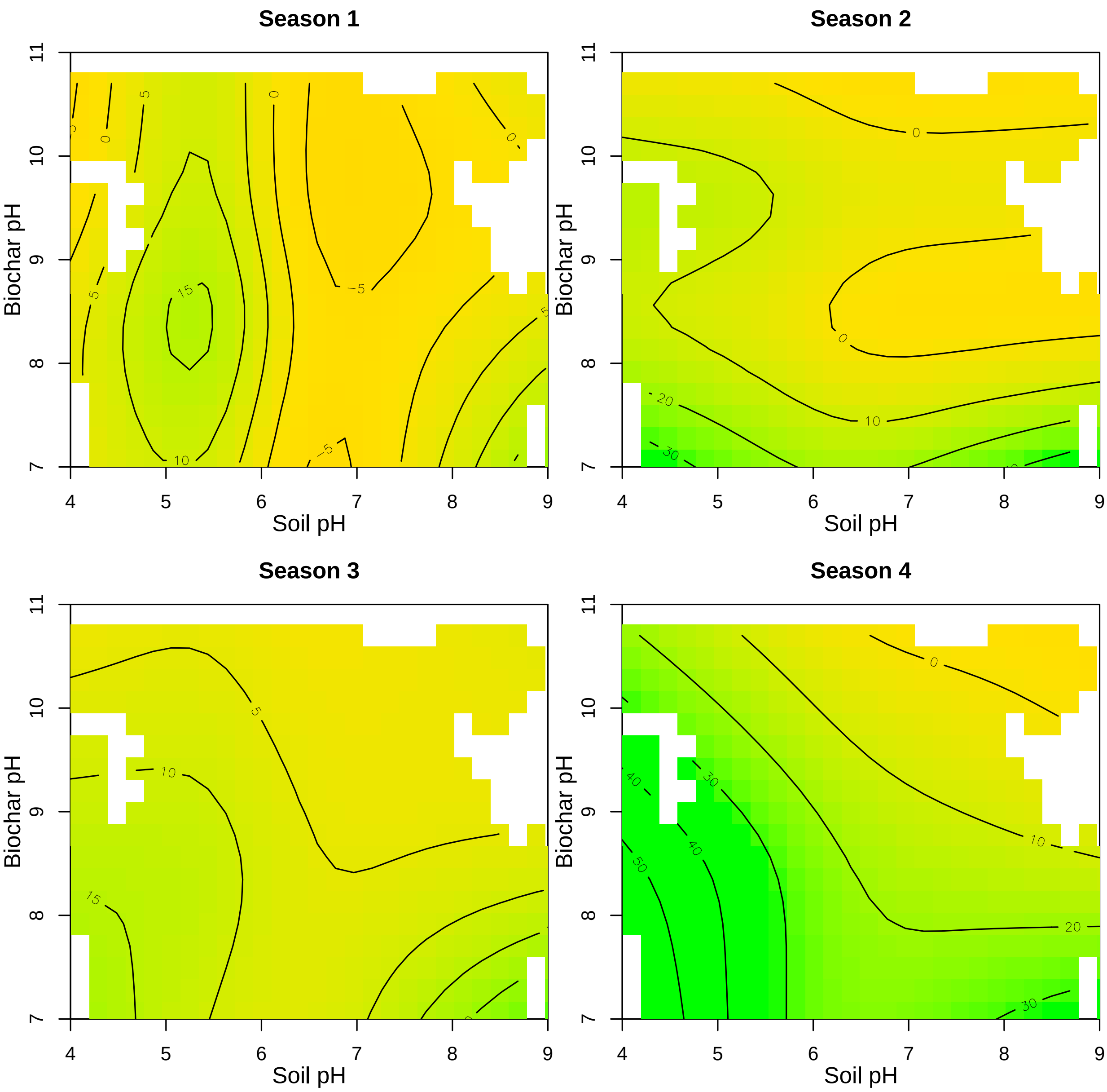


FIGURE A.17: Contour plot giving fitted values of model A.10 over gradients of soil cation exchange capacity, soil organic carbon content, and pH.

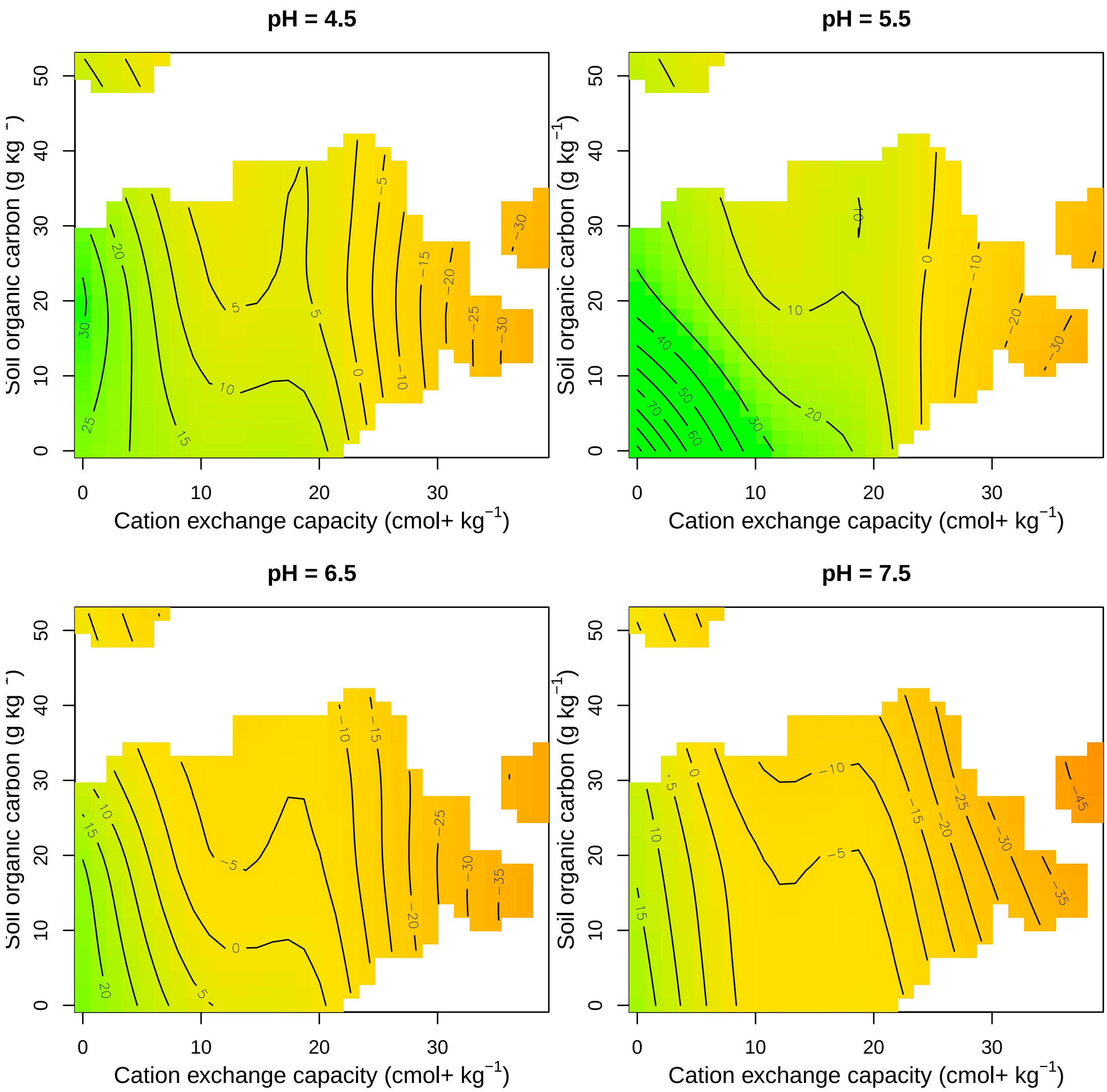
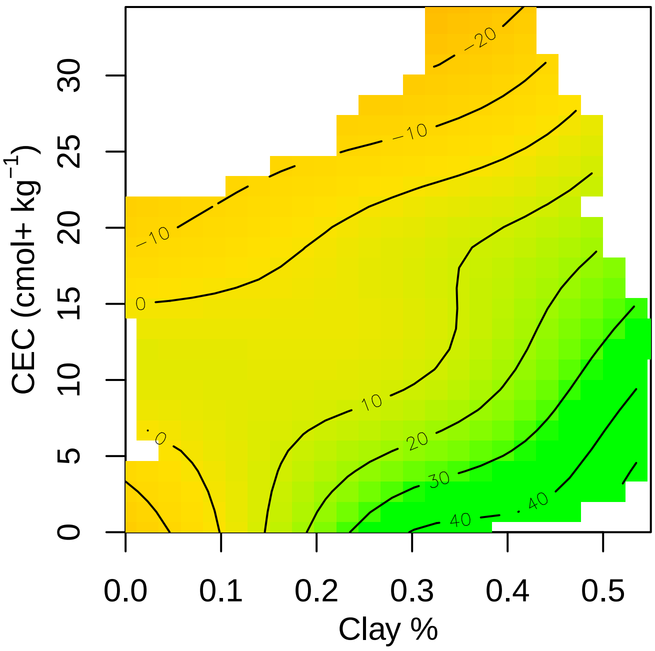


FIGURE A.18: Contour plot giving fitted values of model A.10 over gradients of soil cation exchange capacity and clay content.



at http://andrewcd.berkeley.edu/BCMA_unpublished_studies.zip. The password (required) to decrypt the archive is "biochar."

Figure A.19: Fitted values and confidence intervals for yield response to biochar from model A.10 for non-wood, manure, and wood biochars produced at different temperatures and having different resultant carbon contents.

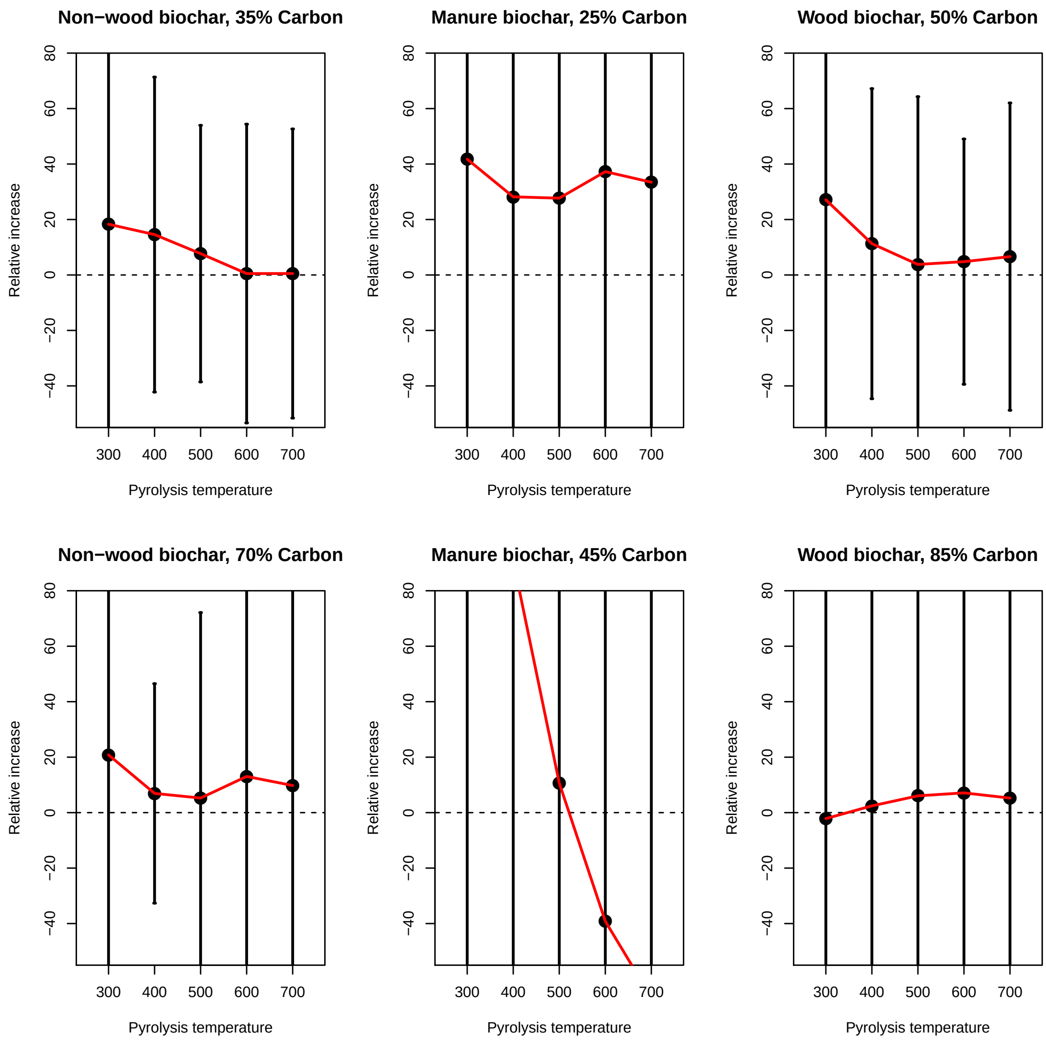
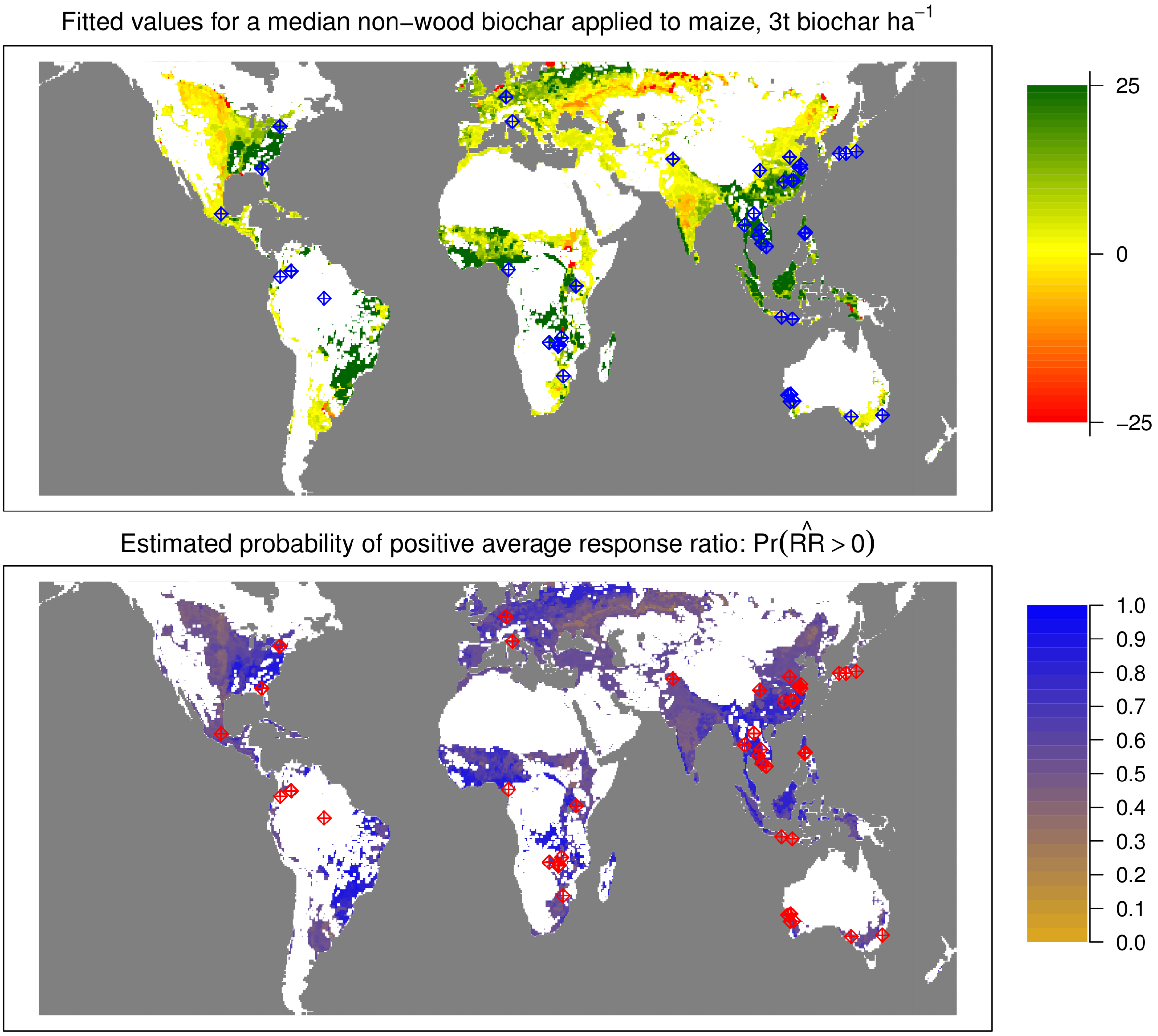


FIGURE A.20: Model A.10 predictions for a median non-wood biochar applied to maize, projected onto spatially-averaged soil property dataset from Batjes (2005)[55]. Upper maps give point estimates of percentage relative increase over control for a .5x.5 degree gridcell, and lower maps give the statistical confidence that the estimated fitted value is greater than zero. Areas with < 5% of land in agriculture are masked using agricultural area extent data from Ramankutty [67].



Appendix B

Biochar profitability and carbon prices: Evidence from Kenya and Vietnam: *Supplementary Information*

Our optimization results are based on the following model:

$$yield_{it} = \alpha_i + P_t + \begin{bmatrix} B_{it} \\ \text{Fert. Exp.}_{it} \\ \text{Fert. Exp.}_{it} \times B_{it} \end{bmatrix}' \Big[\boldsymbol{\beta} + \boldsymbol{\beta}_i \Big] + \epsilon_{it} \tag{B.1}$$

Maize yields are represented as a continuous variable, though in reality they are strictly non-negative. As such, fitted values may fall outside of the feasible range of the outcome variable. This is the case for model 3.1, as is demonstrated in figure B.1. (In addition, we see clear evidence of heteroskedasticity, though the effect of this on inference is accounted for via robust covariance matrix estimation). In addition, we implicitly assume that the marginal response to fertilizer – both for biochar users and non-users – is constant, though it may not be in reality. This violation of model assumptions might lead to biased estimates, so we test for their importance via the following robustness checks:

- We respecify all continuous variables as penalized splines, using generalized additive models[66, 205].
- We add a small constant to all values maize yield before taking their logarithm. We demonstrate that this measure, when modeled as an outcome of biochar and fertilizer use, is inacceptably sensitive to the value of the chosen constant.

- We change the units of the outcome variable from tons per hectare to 10kg units per hectare (i.e.: we multiply by 100), and round to the nearest integer. This yields a discrete measure, that can be modeled assuming a negative binomial distribution[1]. We fit each of these models, both parametrically and nonparameterically. We find that the best-fitting of these models is a nonparametric negative binomial model – which gives results that are qualitatively similar to model 3.1.

B.0.1 Nonparametric specification

We begin by specifying model 3.1 as an additive mixed model, with smoothness selection by restricted maximum likelihood. Results of doing so are given in figure B.2. This specification shows increasing marginal response of yield to fertilizer in the presence of biochar at high values of fertilizer, as well as narrower distributions of random coefficients on the effect of fertilizer. This model has a less favorable AIC value that its parametric counterpart, so we do not favor it. It is likely that the model is unable to distinguish between a convexity in the interaciton between fertilizer and biochar, and an individually-specific idiosynchratic response to fertilizer at high application, given that sparsity of data near at higher fertilizer application rates.

B.0.2 Adding a small constant before logging

Adding a small constant to zero-valued data can facilitate a logarithmic transformation of the data, which in many cases can stabilize variance and help to satisfy other model assumptions. In our case however, we find that our results are unacceptably sensitive to the choice of that constant. We demonstrate this by fitting univariate regressions of our biochar dummy B on $log(yield + c)$, across gradients of c from very small values to 1, and plotting the estimates of $\hat{\beta}$, along with their 95% confidence intervals.

The left plot gives the result of adding c to all yield measurements. Addition of very low values – a common practice – renders our esitmates negative, likely via increasing the leverage of those points on the logarithmic scale. On the other hand, adding 1 to all observations, such that $log(y + c) = 0$ at $y = 0$, renders our estimate positive and statistically significant. The right plot shows the result of adding a small constant only to those values that contain zeros. Similarly, very small values increase leverage and render point estimates negative, while higher values generate positive point estimates.

Despite this model's lack of robustness to the choice of added constant, we fit it anyway (choosing c=0.1) and present the result in figure B.4. Results are qualitatively similar

[1]We use a negative binomial rather than a poisson, as poisson is not scale invariant and not robust to overdispersion. Negative binomial, on the other hand, accounts for overdispersion with respect to the poisson by virtue of the fact that a dispersion parameter is estimated by standard negative binomial regression software.

to the those given in figure 3.1, though it is apparent that the variance of the random coeffients on the effect of biochar have increased.

The logged model is sensitive to the assumption of linearity in its functional form – respecifying fertilizer and its interaction with biochar nonparametrically using a generalized additive model improves AIC, and suggests a locally negative slope in biochar's interaction with fertilizer, before flattening and rising at higher values, and becoming statistically insignificant B.5.

B.0.3 Negative binomial models

Allison & Waterman [206] provide a simulation study demonstrating that negative binomial regression is robust to the incidental parameters problem, which can bias the estimates of nonlinear models with many fixed effects. While our data is continuous, we change its scale before rounding, to achieve a smoother discrete distribution. Unlike poisson, the negative binomial distribution estimates two parameters in the course of its fitting – a mean as well as a dispersion parameter – rendering it less scale-dependent than a poisson. Specifically, we multiply our yield variable by 100 before rounding, and fit a log-linked model with the linear predictor specified identically to that of the gaussian models previously considered – both parametrically and nonparametrically. Results are given in figures B.6 and B.7. Of these two, the nonparametric model has better AIC. Results are qualitatively similar to the logged model, which in turn are similar to the linear parametric specification.

FIGURE B.1: Diagnostic plots for model B.1.

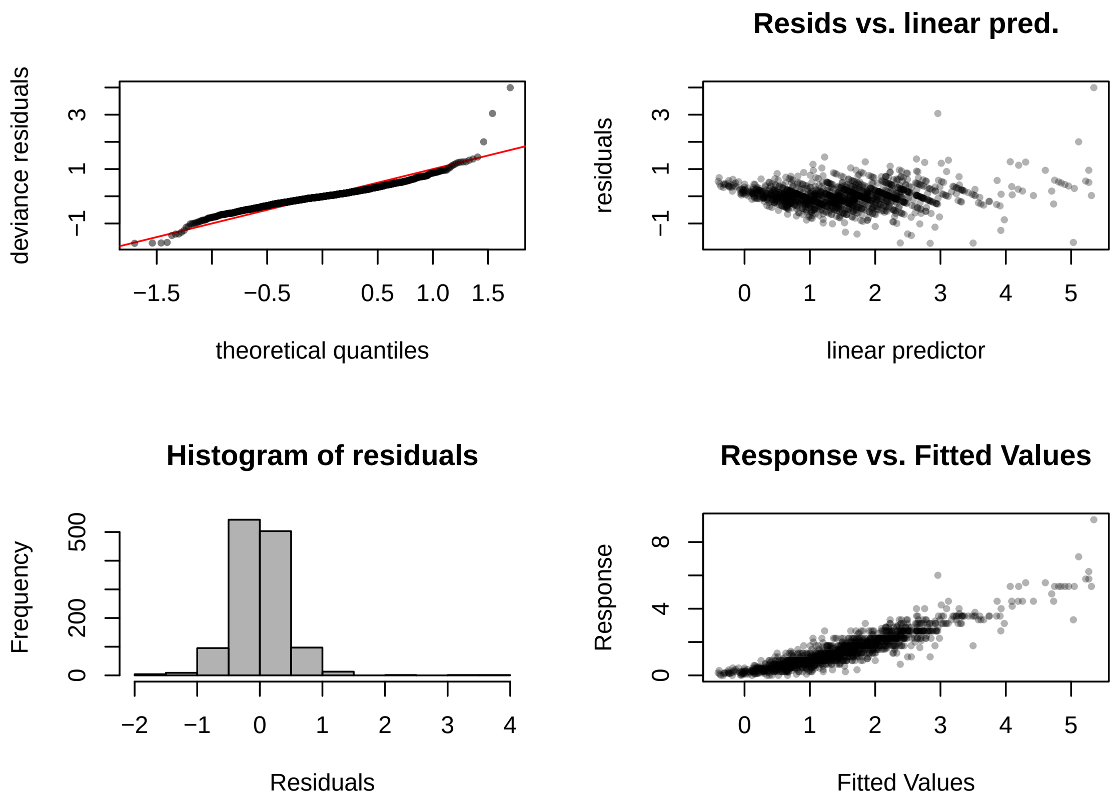


FIGURE B.2: Estimates of model 3.1 where the continuous variables have been represented by penalized splines.

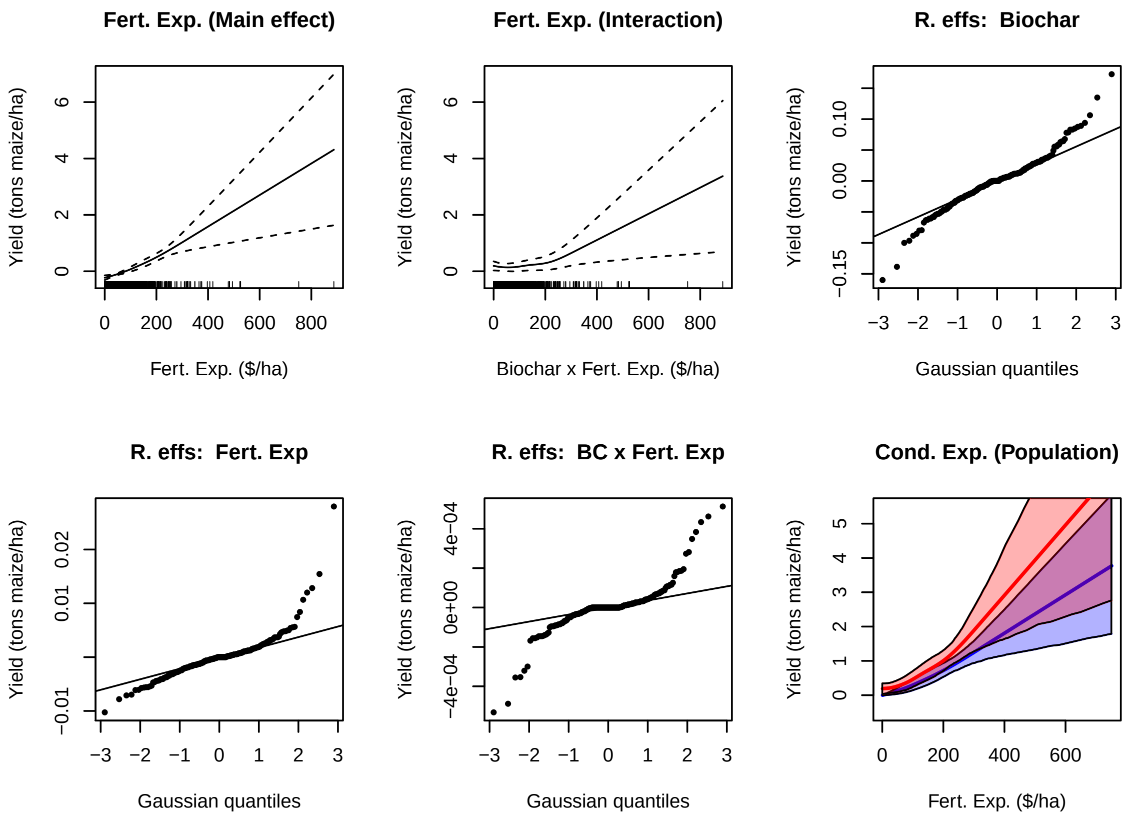


FIGURE B.3: Estimates of β from models of the form $log(y + c) = \alpha + \beta B + \epsilon$, across gradients of c.

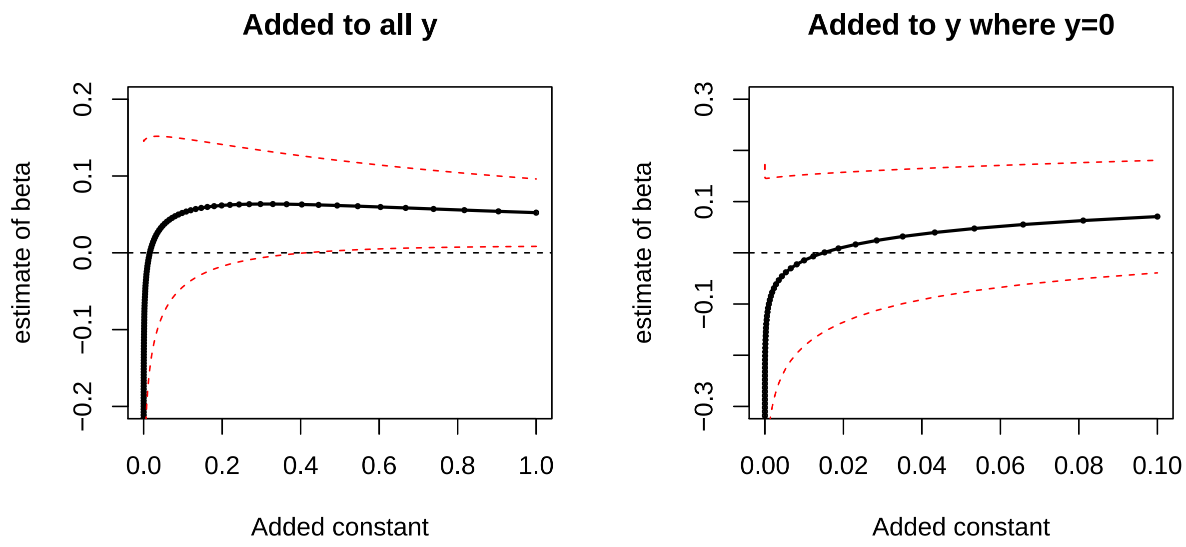


Figure B.4: Estimates of model 3.1 where the outcome variable has been respecified as $log(yield + .01)$.

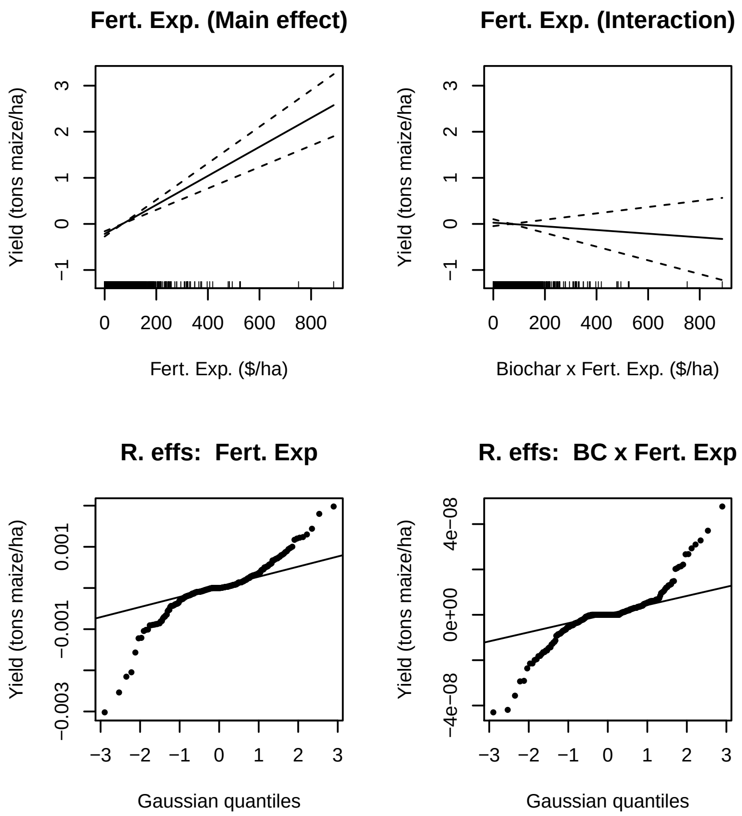


Figure B.5: Estimates of model 3.1 where the outcome variable has been respecified as $log(yield + .01)$ and the continuous covariates havebeen represented by penalized splines.

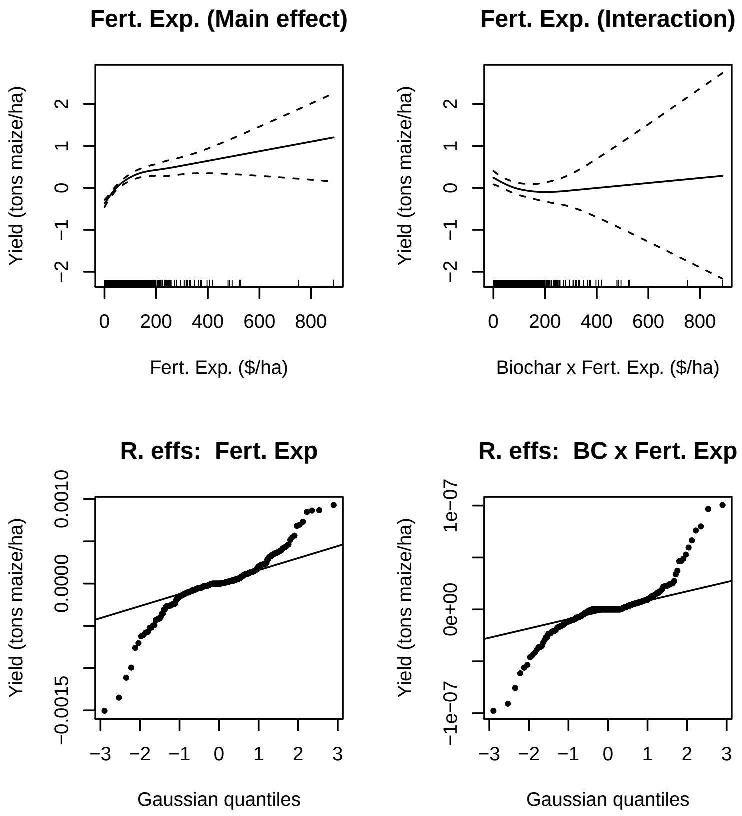


FIGURE B.6: Estimates of model 3.1 where the outcome variable has been multiplied by 100 and rounded, and specified to follow a negative binomial distribution.

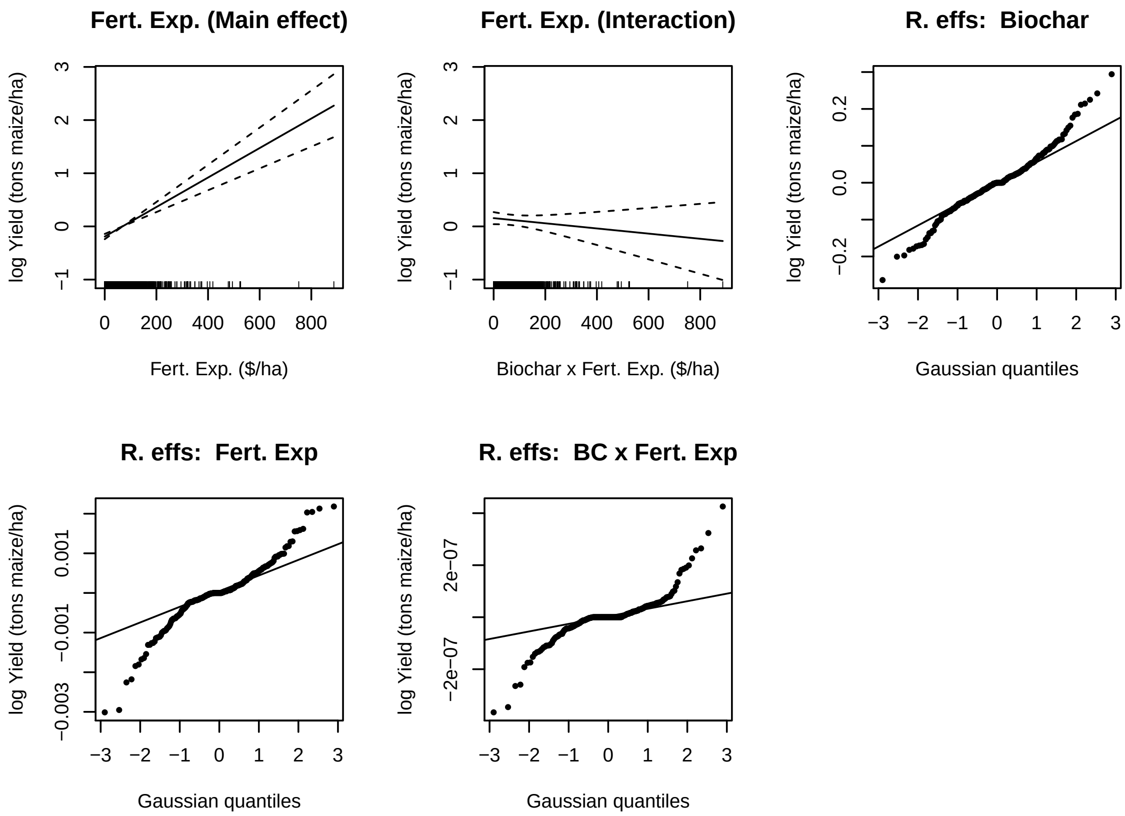
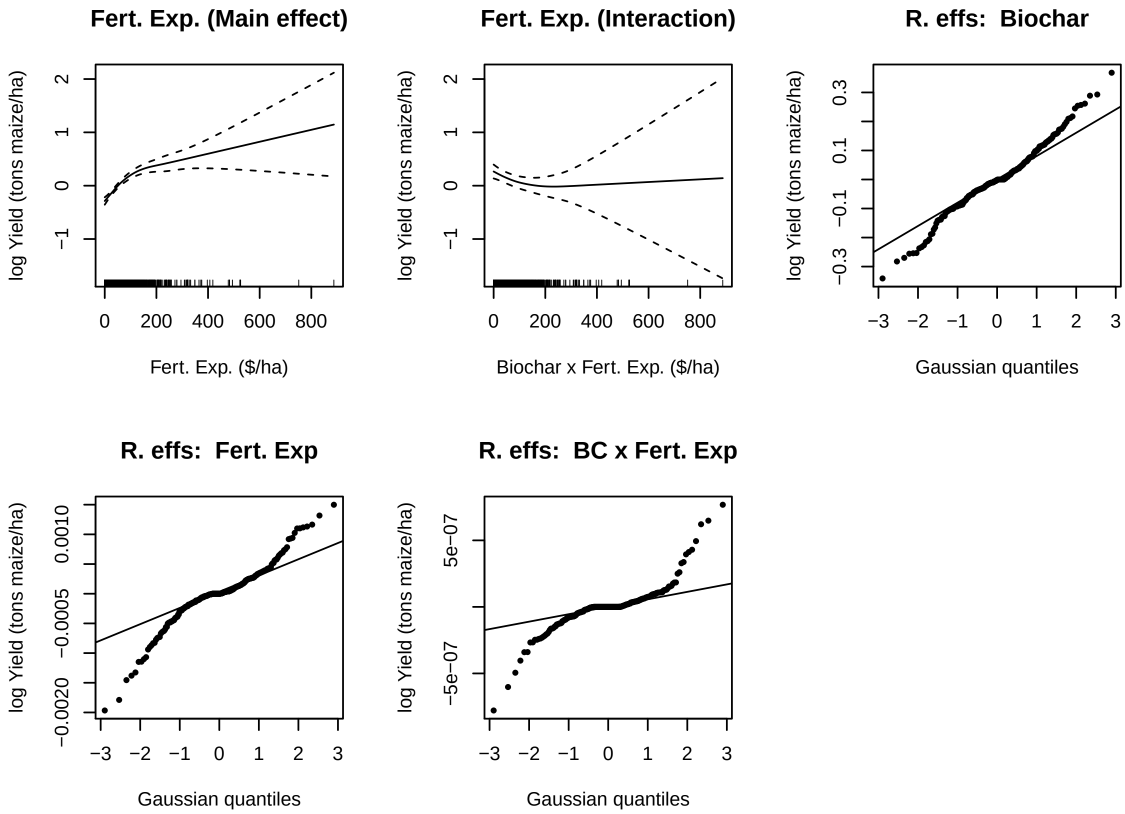


Figure B.7: Estimates of model 3.1 where the outcome variable has been multiplied by 100 and rounded, and specified to follow a negative binomial distribution. Continuous covariates are represented by penalized splines.



Appendix C

Subsidies, Social Learning, and Charcoal Dust: Impact and Adoption Dynamics of Biochar in Western Kenya: *Supplementary Information*

Table C.1: Comparison of summary statistics by demonstration plot status; taking a demonstration plot, not being offered a demonstration plot, and refusing a demonstration plot. The sixth and seventh columns respectively give the p values of the differences between those with and without demonstration plots, and the differences between those with, without, and refusing. The latter are p-values associated with F-statistics.

	All	1 Got demo	2 No demo	3 Refused demo	p(diff(1,2))	p(diff(1,2,3))
N	1115	72	1034	9	-	-
Gender (male=1)	0.25	0.28	0.24	0.11	0.53	0.53
Maize yield (t/ha)	1.34	1.30	1.34	1.79	0.71	0.25
Maize hectares	0.35	0.31	0.35	0.48	0.28	0.28
Beans yield	0.23	0.25	0.23	0.10	0.58	0.51
Beans hectares	0.33	0.29	0.33	0.76	0.34	0.02*
Fert. exp. per hectare	88.97	87.40	88.95	103.98	0.86	0.82
Acres sugar	0.10	0.11	0.10	0.01	0.72	0.62
Number listing as social contact	92.73	93.94	92.59	99.13	0.83	0.91
Eigenvector centrality (inbound)	0.21	0.25	0.21	0.23	0.11	0.26
Distance from center of zone	0.01	0.01	0.01	0.01	0.47	0.76
Owns bicycle	0.62	0.46	0.63	1.00	0.09*	0.09*
Number of phones in household	1.11	1.08	1.11	1.33	0.78	0.74
Income per year at baseline	346.55	400.22	341.98	484.71	0.61	0.80
Owns motorcycle	0.07	0.10	0.07	0.22	0.49	0.35
Heard of Re:char	0.04	0.03	0.04	0.11	0.61	0.48
Heard of ACON	0.09	0.11	0.09	0.11	0.59	0.85
Heard of the Rutuba	0.05	0.10	0.05	0.11	0.09*	0.17
Has used charcoal as fertilizer	0.05	0.06	0.05	0.11	0.69	0.61
Has produced charcoal	0.33	0.40	0.32	0.44	0.14	0.26

Figure C.1: Locations of the four zones in which the project operated, in relation to each other and Bungoma town. Dots represent households in the study. Inset map shows Bungoma's location in Kenya.

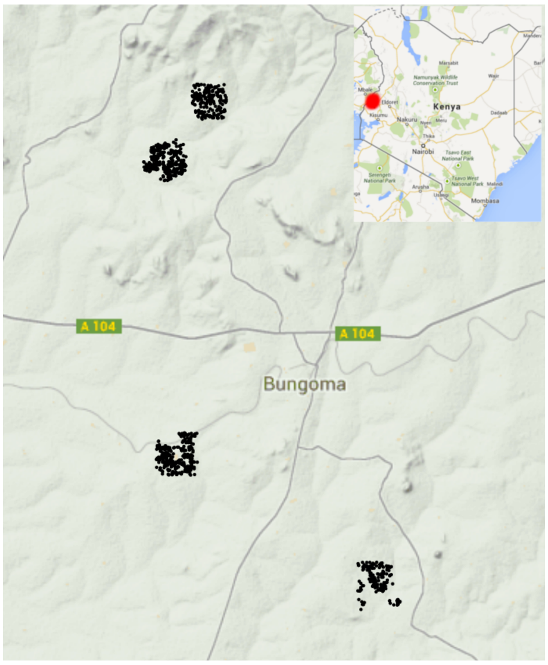


Figure C.2: Signs placed on biochar demonstration plots. "Bila" means "without" in Kiswahili.

Figure C.3: The spatial distribution of the biochar demonstration plots, established in August and September 2013. Each zone is approximately 4km^2.

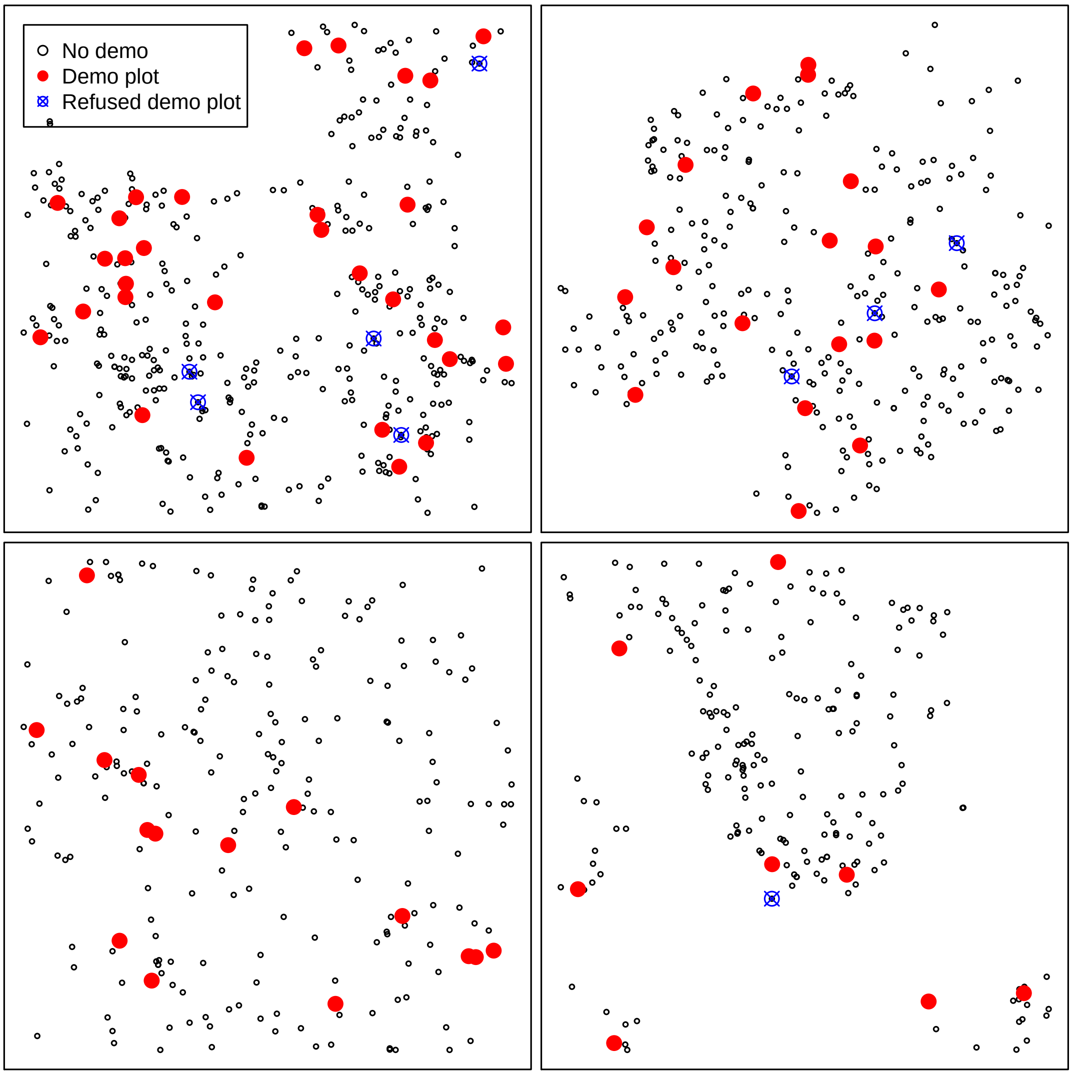


Table C.2: Comparison of summary statistics by farmers who attrited after the baseline survey, versus those who didn't. The last column gives the p-value of the difference between the two.

	All	Full Sample	Attrited	p
N	1115	826	289	-
Gender	0.25	0.27	0.16	$<.01^{***}$
Number listing as social contact	92.73	94.01	89.07	0.17
Eigenvector centrality (inbound)	0.21	0.21	0.22	0.54
Distance from center of zone	0.01	0.01	0.01	0.70
Owns bicycle	0.62	0.61	0.65	0.42
Number of phones in household	1.11	1.13	1.06	0.23
Income per year at baseline	346.55	321.24	420.77	0.11
Owns motorcycle	0.07	0.08	0.06	0.51
Hear of Re:char	0.04	0.04	0.04	0.83
Heard of ACON	0.09	0.09	0.09	0.82
Hear of the Rutuba	0.05	0.06	0.05	0.64
Has used charcoal as fertilizer	0.05	0.05	0.03	0.26
Has produced charcoal	0.33	0.33	0.31	0.55

TABLE C.3: Results of model 4.1, relating probability of linkage to observed covariates including distance, attendance of children in common schools, attendance of common churches, and gender, for different dimensions of social network linkage.

			In past month...			Know from...		
	Linkage	Talk farming	Seen	Seen land	Seen crops	Is a relative	Church	Savings group
(Intercept)	-10.67^{***}	-11.17^{***}	-10.65^{***}	-11.54^{***}	-11.38^{***}	-9.59^{***}	-7.87^{***}	-8.69^{***}
	(0.10)	(0.14)	(0.10)	(0.12)	(0.11)	(0.11)	(0.12)	(0.17)
In same church	0.51^{***}	0.75^{***}	0.52^{***}	0.71^{***}	0.58^{***}	0.68^{***}	2.80^{***}	1.39^{***}
	(0.04)	(0.06)	(0.04)	(0.05)	(0.05)	(0.05)	(0.05)	(0.07)
Kids in same school	0.20^{***}	0.18^{***}	0.20^{***}	0.15^{***}	0.15^{***}	0.18^{***}	0.23^{***}	0.21^{**}
	(0.03)	(0.05)	(0.03)	(0.04)	(0.04)	(0.04)	(0.05)	(0.07)
Gender_j	0.16^{*}	0.09	0.18^{*}	0.18^{**}	0.20^{**}	0.26^{***}	0.09	-0.02
	(0.08)	(0.07)	(0.07)	(0.06)	(0.07)	(0.07)	(0.07)	(0.08)
Gender_i	0.54^{***}	0.94^{***}	0.55^{***}	0.74^{***}	0.71^{***}	0.45^{***}	0.03	-0.28
	(0.08)	(0.14)	(0.08)	(0.09)	(0.10)	(0.09)	(0.08)	(0.17)
$\text{Gender}_i == \text{Gender}_j$	0.08^{**}	0.18^{***}	0.07^{*}	0.12^{***}	0.09^{**}	0.03	0.01	0.07
	(0.03)	(0.04)	(0.03)	(0.03)	(0.03)	(0.04)	(0.05)	(0.06)
$\log(\text{distance}_{i,j})$	-1.86^{***}	-1.34^{***}	-1.80^{***}	-1.73^{***}	-1.76^{***}	-1.21^{***}	-0.75^{***}	-0.67^{***}
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)
Log Likelihood	-27709.87	-13885.86	-27162.71	-21695.06	-23069.07	-17344.47	-12882.75	-7861.72
Deviance	55419.74	27771.73	54325.42	43390.12	46138.14	34688.95	25765.51	15723.44
Num. obs.	70950	71025	71025	71025	71025	71025	71025	71025
Num. groups: Asked about	1116	1116	1116	1116	1116	1116	1116	1116
Num. groups: Asked	946	947	947	947	947	947	947	947
Variance: $\text{ID}_j \times \log(\text{dist})$	0.05	0.02	0.04	0.02	0.03	0.03	0.02	0.02
Variance: $\text{ID}_i \times \log(\text{dist})$	0.04	0.12	0.04	0.06	0.07	0.05	0.03	0.12

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

FIGURE C.4: Distributions of absolute prediction error from estimates of 4.1 across multiple social network dimensions. Prediction error defined as $|\mathbf{1}(link_{ij}) - \widehat{pr}(link_{ij})|$. Dashed red lines represent means.

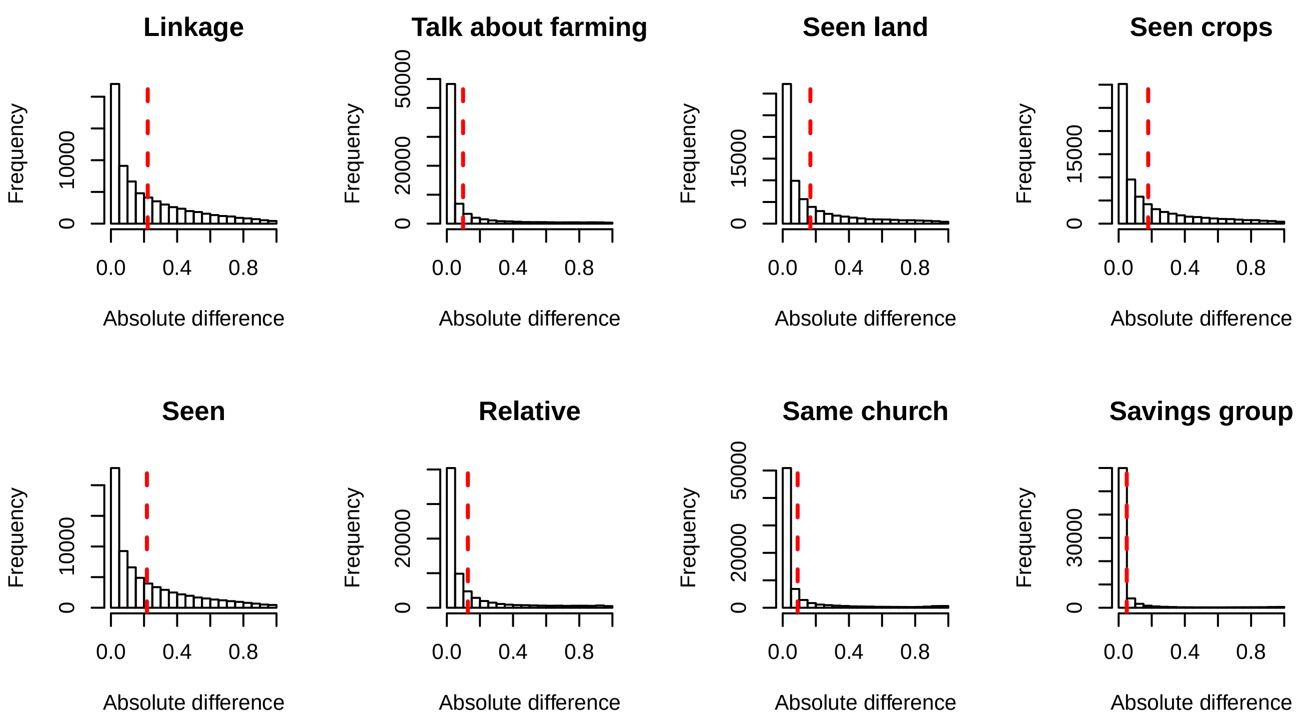


FIGURE C.5: Theoretical utility of biochar adoption at a cost of \$125/ha (KSh850 per quarter acre plot), for a hypothetical decisionmaker with a mean-standard deviation utility function of the form $U(\mu, \sigma) = \mu^\theta - \sigma^\gamma$. The mean and standard deviation of observed profits from demonstration plots are \$27 and \$140 per hectare, respectively. The bold contour line represents the combination of taste parameters for the mean and standard deviation at which the modeled decisionmaker would be indifferent to biochar adoption. Following Saha [168], $\gamma > 0$ corresponds to risk aversion, and $\theta < 1$ corresponds to declining average risk aversion (conditional on $\gamma > 0$.) This plot demonstrates that certain sorts of risk aversion could help explain the attractiveness (or lack thereof) of biochar adoption.

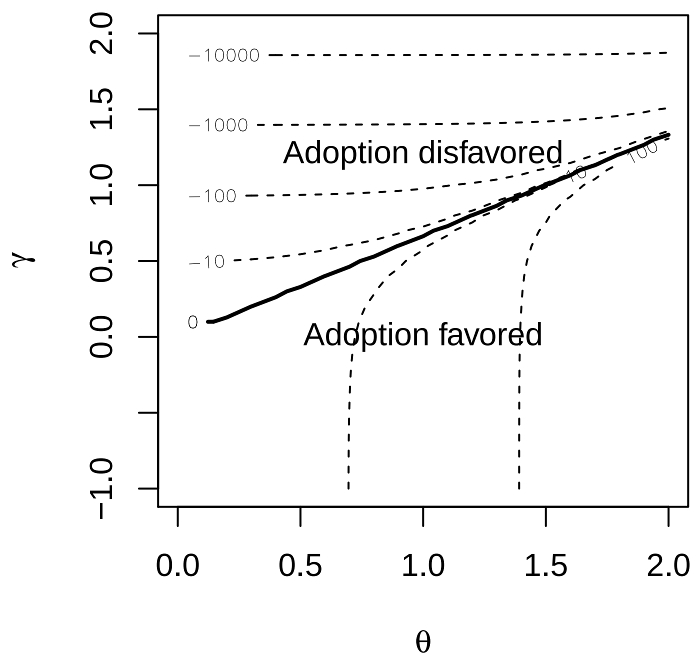
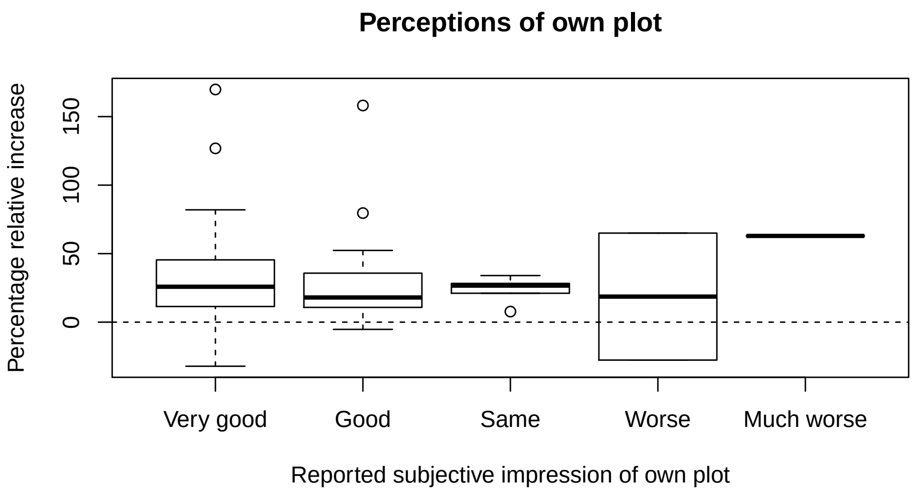


Figure C.6: Left: Likert-scale responses to the question "how were the yields on the biochar portion of the demonstration plot, compared to the portion of the demonstration plot without biochar." Right: Likert scale responses to the question "among others you know with biochar demonstration plots, how were yields on the biochar portion of their demonstration plots, compared to the portions of their demonstration plots without biochar."

Table C.4: Average measured demonstration plot relative increases among social network contacts, by stated perception of the relative magnitude of the average differences, across different dimensions of social network linkage. We include both reported (sampled) networks (upper panel) and probabilistically imputed networks (lower panel).

		Very Good	Good	Same	Didn't See
Average demo plot relative increase among reported...	Links	25.53	28.36	31.53	28.23
	Links with whom farming discussed	24.45	27.95	26.90	24.05
	Links seen in past month	27.71	29.25	28.66	28.66
	Links seen crops in past month	26.20	27.77	32.46	26.64
	Links seen land in past month	25.24	26.74	28.27	28.73
	Relatives	25.41	25.52	58.55	30.99
	Fellow church members	25.64	28.02	28.13	23.70
	Fellow savings group members	26.72	34.21	80.84	30.41
Average demo plot relative increase among imputed...	Links	27.32	28.14	30.18	28.04
	Links with whom farming discussed	25.92	26.35	26.14	25.63
	Links seen in past month	28.75	29.30	29.89	29.39
	Links seen crops in past month	27.20	27.47	29.49	27.34
	Links seen land in past month	26.76	26.75	28.19	27.67
	Relatives	28.45	28.64	39.05	30.10
	Fellow church members	25.98	26.71	27.38	26.21
	Fellow savings group members	29.16	29.16	33.34	29.09

Figure C.7: Upper: The distribution of measured percentage differences in yield between biochar and control portions of demonstration plots, by stated perception of the relative magnitude of the difference, for holders of demonstration plots. Lower: the distribution of weighted average measured percentage differences in yield between biochar and control portions of demonstration plots within (imputed) social networks, by stated perception of the relative magnitude of the average differences, for the full sample.



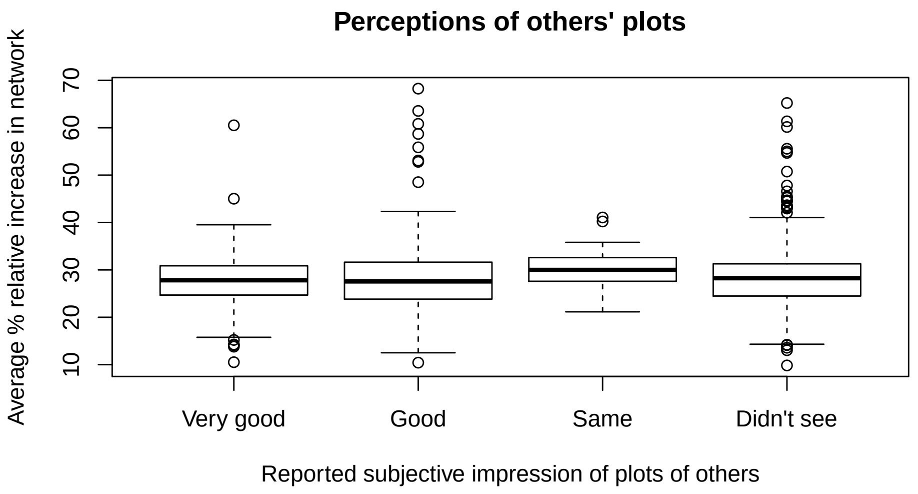


FIGURE C.8: Univariate logistic GAMs relating uptake in early 2014 to the share of each farmer's network adopting biochar (left), the number of links in each farmers network (right), and biochar price (bottom). The right logit controls for the total number of links in each farmer's network, and the fit represents the marginal effect at the mean number of network links.

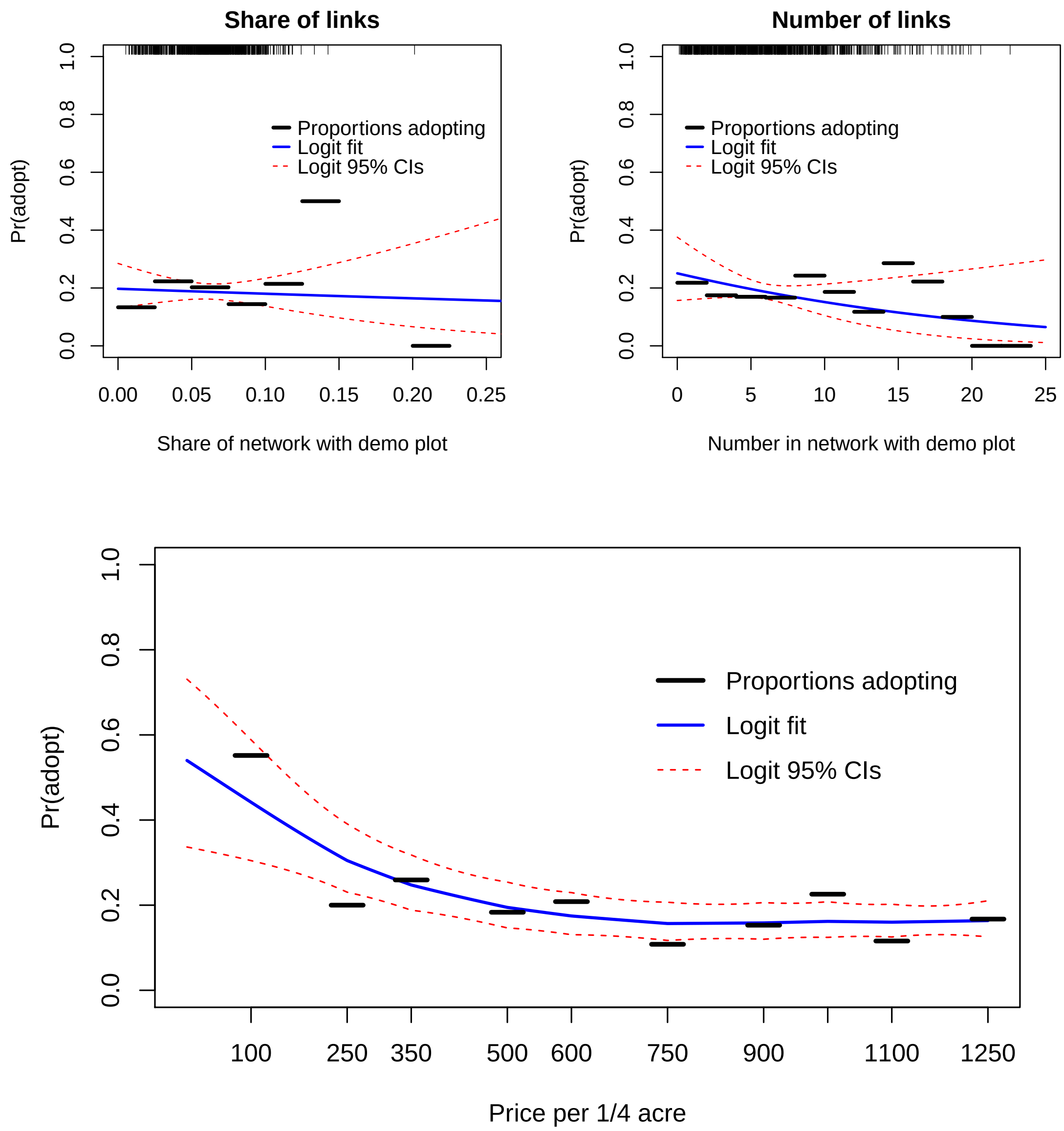


FIGURE C.9: The distribution of biochar prices and risk-free trials offered in the 2014 long rains season.

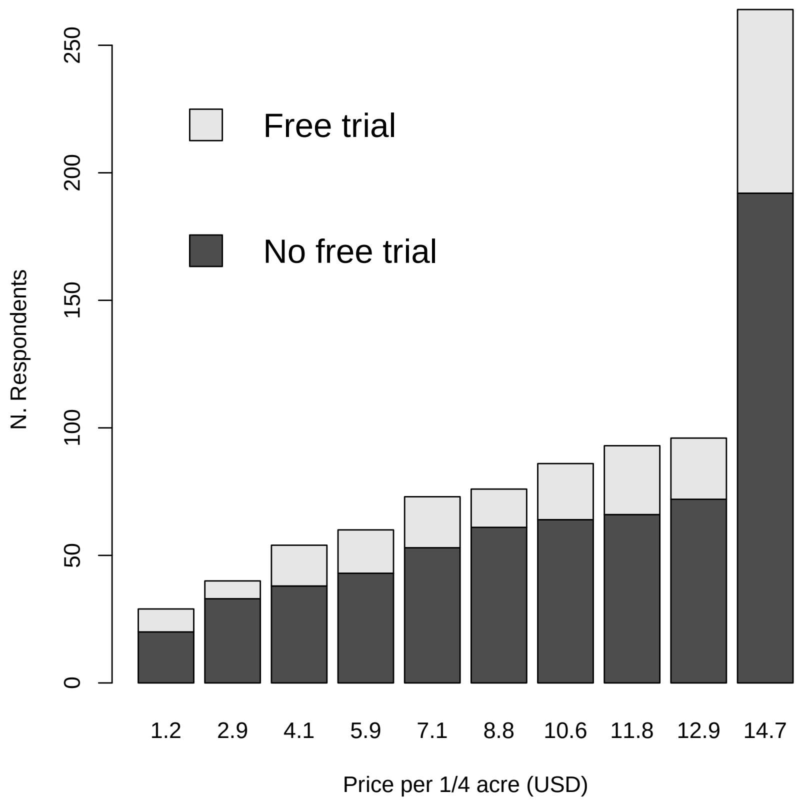


Table C.5: Logistic regression coefficients and naive p-values relating biochar uptake to the shares (left) and numbers (right) of various dimensions of a social network taking up biochar, using observed ("Obs.") and imputed ("impd") networks. None of the significant p-values are robust to a Holm-Bonferonni correction.

		Share of network			Number in network		
		$\hat{\beta}$	$se(\hat{\beta})$	$p(\lvert\hat{\beta}\rvert) > 0$	$\hat{\beta}$	$se(\hat{\beta})$	$p(\lvert\hat{\beta}\rvert) > 0$
Links	Obs.	0.42	1.35	0.76	-0.02	0.07	0.79
	Impd.	-1.00	3.77	0.79	-0.06	0.05	0.20
Talked farming	Obs.	-0.27	0.75	0.72	0.01	0.13	0.94
	Impd.	-0.76	1.90	0.69	-0.03	0.10	0.79
Seen land	Obs.	0.79	0.95	0.40	0.02	0.09	0.80
	Impd.	1.74	2.45	0.48	-0.04	0.07	0.60
Seen Crops	Obs.	0.08	0.99	0.94	-0.03	0.09	0.75
	Impd.	0.08	2.76	0.98	-0.07	0.07	0.31
Seen	Obs.	0.33	1.31	0.80	-0.04	0.07	0.54
	Impd.	-0.49	3.57	0.89	-0.08	0.05	0.11
Relatives	Obs.	0.17	0.70	0.81	-0.06	0.13	0.65
	Impd.	-0.70	2.47	0.78	-0.09	0.11	0.40
Savings Group	Obs.	1.22	0.59	0.04	0.23	0.20	0.25
	Impd.	3.60	2.07	0.08	0.17	0.18	0.34
Church	Obs.	1.24	0.53	0.02	0.24	0.15	0.11
	Impd.	5.76	2.30	0.01	0.18	0.13	0.16
Distance	Within 1km	-1.71	3.83	0.65	-0.01	0.03	0.75
	Within 2km	-3.83	5.36	0.48	0.02	0.03	0.44
EV cent. weighted	Impd.	-5.51	11.87	0.64	-0.06	0.16	0.71

Table C.6: Parametric coefficient estimates and goodness-of-fit statistics from various specifications of model 4.3. Nonparametric components of these models are given in figure C.10. The dependent variable is adoption of biochar in early 2014 (in the runup to the long rains season), and all coefficients represent marginal effects on the logit scale.

	Spec. 1	Spec. 2	Spec. 3	Spec. 4	Spec. 5	Spec. 6	Spec. 7
Intercept	-4.39^{***}	-5.47^{***}	-5.46^{***}	-5.45^{***}	-5.59^{***}	-5.46^{***}	-5.44^{***}
	(0.47)	(0.79)	(0.78)	(0.78)	(0.82)	(0.78)	(0.78)
Gender	-0.19	-0.14	-0.10	-0.15	-0.14	-0.14	-0.16
	(0.32)	(0.32)	(0.32)	(0.32)	(0.32)	(0.32)	(0.32)
Mateka (zone)	-0.47	-0.47	-0.50	-0.45	-0.43	-0.46	-0.48
	(0.42)	(0.42)	(0.42)	(0.42)	(0.42)	(0.42)	(0.42)
Mukwa (zone)	0.09	0.04	0.03	0.05	0.09	0.05	0.05
	(0.42)	(0.42)	(0.42)	(0.42)	(0.42)	(0.42)	(0.42)
Nasianda (zone)	0.09	0.06	0.05	0.09	0.08	0.06	0.05
	(0.49)	(0.49)	(0.50)	(0.50)	(0.49)	(0.49)	(0.49)
Heard of Re:char	-0.28	-0.21	-0.20	-0.20	-0.21	-0.21	-0.22
	(0.66)	(0.63)	(0.64)	(0.64)	(0.64)	(0.63)	(0.64)
Heard of ACON	1.56^{**}	1.60^{**}	1.54^{**}	1.59^{**}	1.62^{**}	1.60^{**}	1.59^{**}
	(0.54)	(0.56)	(0.56)	(0.56)	(0.56)	(0.56)	(0.56)
Ever used char as fert.	-0.77	-0.86	-0.85	-0.89	-0.84	-0.85	-0.87
	(0.73)	(0.74)	(0.75)	(0.75)	(0.74)	(0.74)	(0.73)
Heard of Rutuba	-0.43	-0.29	-0.27	-0.32	-0.31	-0.29	-0.29
	(0.63)	(0.61)	(0.61)	(0.61)	(0.61)	(0.62)	(0.61)
Got demo plot	1.30^{*}	1.39^{*}	1.36^{*}	1.40^{*}	1.84^{*}	1.36^{*}	1.21
	(0.55)	(0.57)	(0.56)	(0.56)	(0.80)	(0.67)	(0.62)
Offered RFT	5.11^{***}	6.09^{***}	6.09^{***}	6.06^{***}	6.19^{***}	6.08^{***}	6.06^{***}
	(0.40)	(0.73)	(0.73)	(0.72)	(0.76)	(0.73)	(0.73)
Demo x RFT					-0.81		
					(1.05)		
AIC	415.11	407.48	408.59	408.49	408.91	409.48	408.64
BIC	509.04	501.32	514.88	508.10	507.59	508.14	508.43
Deviance explained	0.55	0.56	0.56	0.56	0.56	0.56	0.56
Num. obs.	863	863	863	863	863	863	863
Num. smooth terms	4	5	6	6	5	6	6

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Figure C.10: Estimates of nonparametric terms from various specifications of 4.3. Estimates of parametric terms, and goodness-of-fit measures given in table C.6. y-axes are common to specifications (rows) and are estimated probability of adoption on the logit scale. Continuous-by-continuous interactions (price and network share with biochar) represented by tensor products of the marginal bases, and the z-axis is probability of adoption on the logit scale, with x and y-axes given in the plot. p-values for nonparametric terms based on reduced-rank Wald tests for joint equality to **0** among all penalized coefficients comprising a term [80]. Acronyms: “DCOZ” = “Distance from the center of the zone,”, “R” = “Risk-Free Trial,”, “DP” = “Demo Plot.”

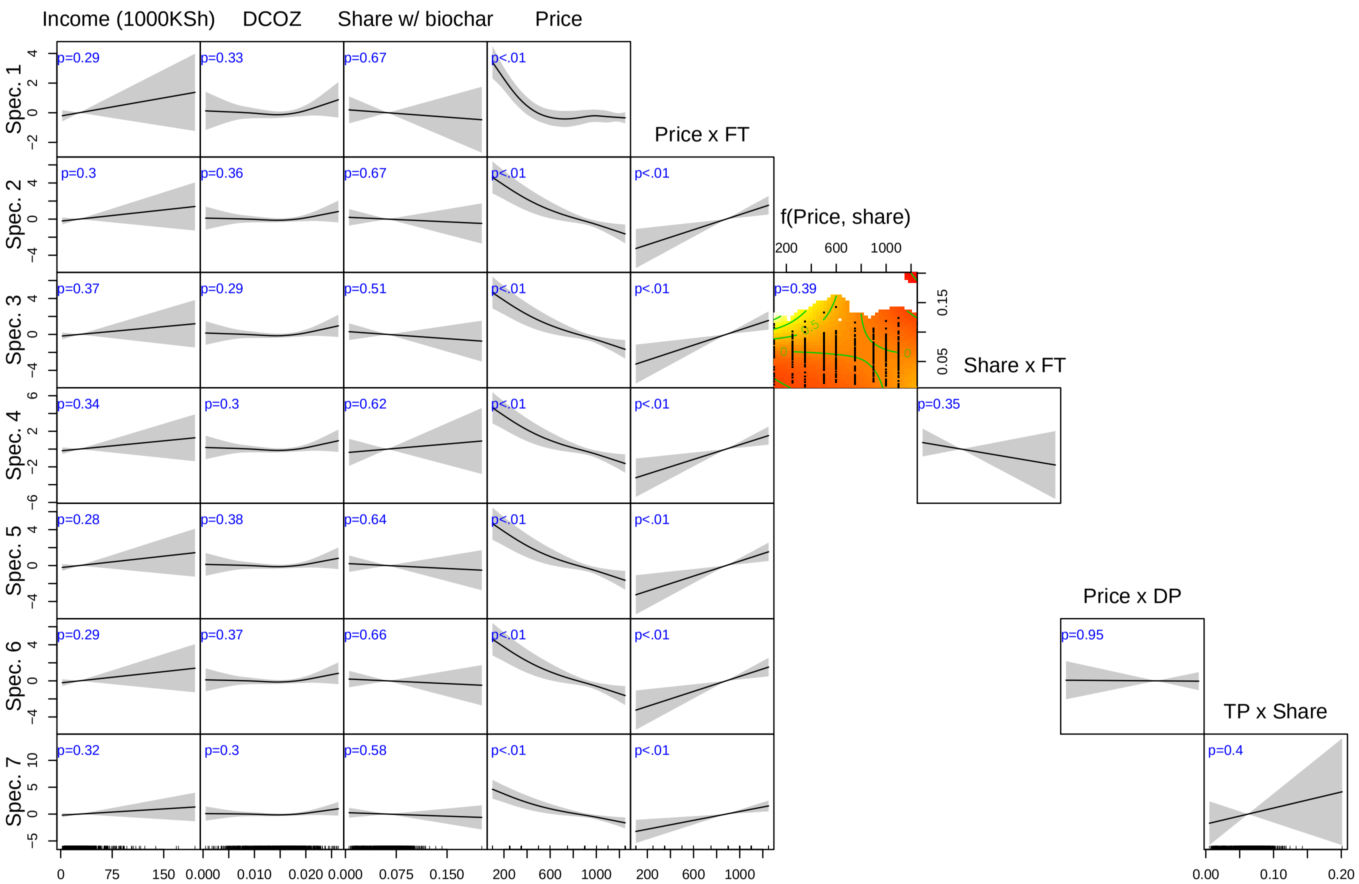
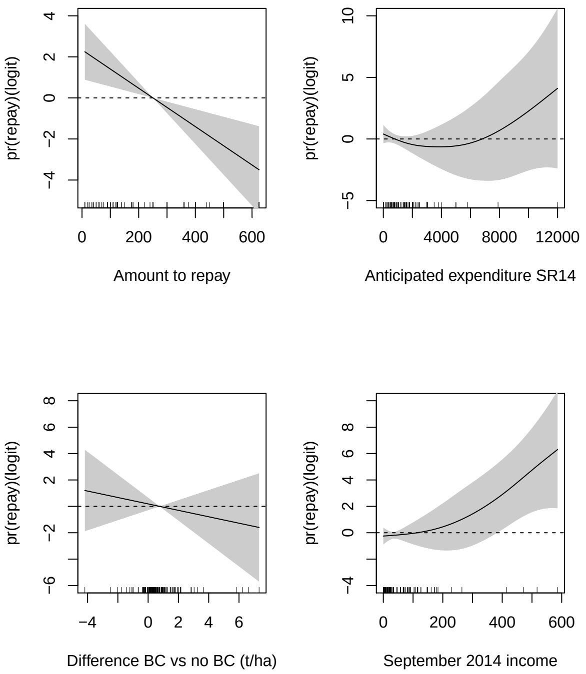


Table C.7: Results of a logistic GAM relating RFT repayment (binary) to whether or not the respondent had already planted, their reported subjective impression of the results of the biochar that they got, the amount that they had to repay, their reported expected expenditure for the 2014 short rains season, the measured difference between biochar yields and control yields, and their reported income in the previous month.

(Intercept)	-4.82^{**}
	(1.59)
Not yet planted	1.76
	(0.91)
Partially planted	0.53
	(0.82)
Impression: "Very good"	0.46
	(1.06)
Impression: "Good"	-1.03
	(1.10)
Impression: "Bad"	0.25
	(1.74)
Impression: "Very bad"	0.64
	(1.21)
Deviance explained	0.32
Num. obs.	132



$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

FIGURE C.11: Comparisons of several specifications of model 4.3 – OLS, additive Gaussian, parametric logit, and additive Gaussian. Marginal effects at means of non-plotted covariates given for Gaussian models (a) and logit model (b). Comparison of predictions for different models given in (c). Differences in marginal effects at means of covariates given in (d) and (e) by RFT status.

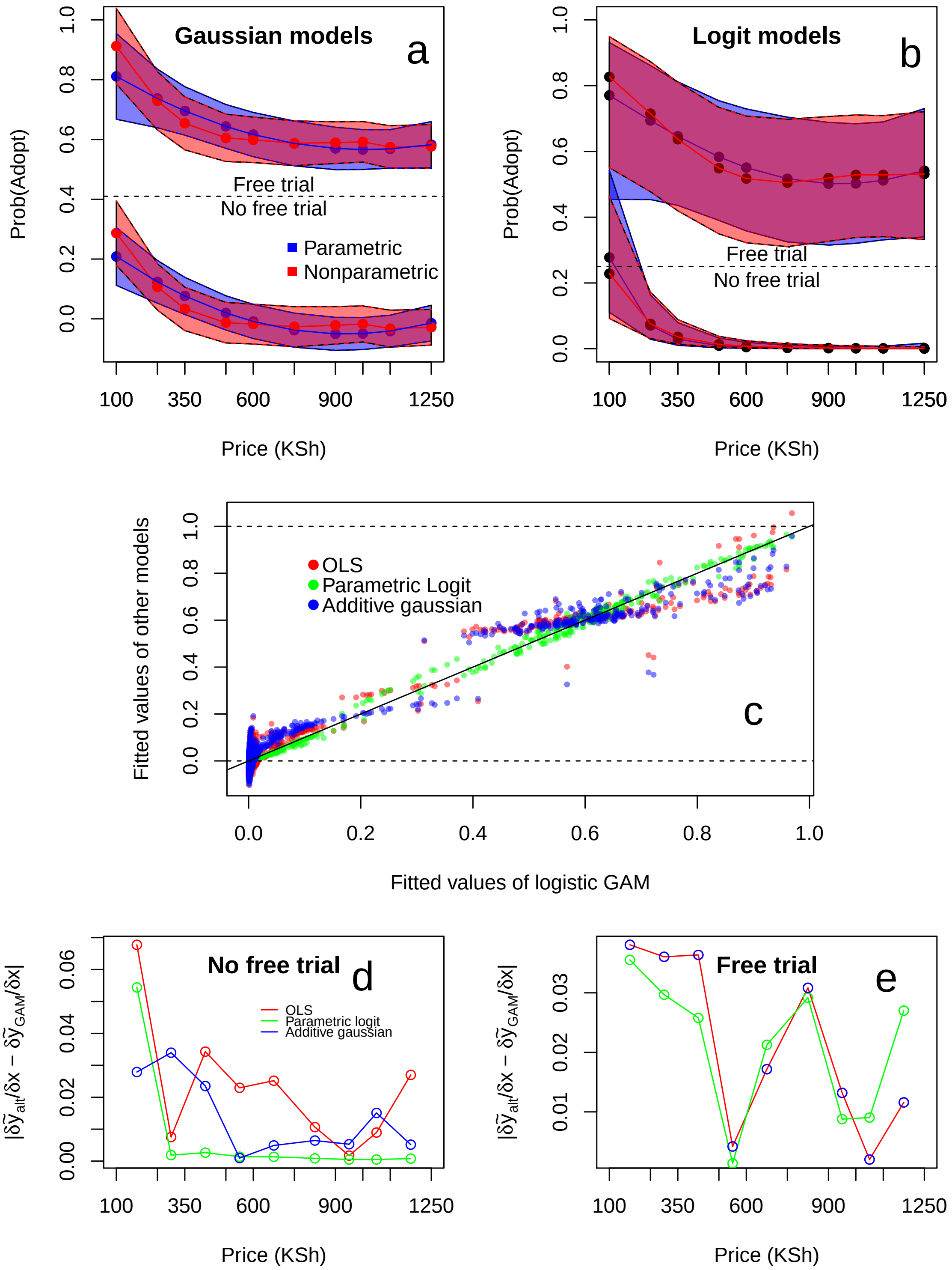


FIGURE C.12: Univariate nonparametric logit fits relating the average price for (partially-imputed) network links (upper) and the proportion of the network having a RFT offer (lower).

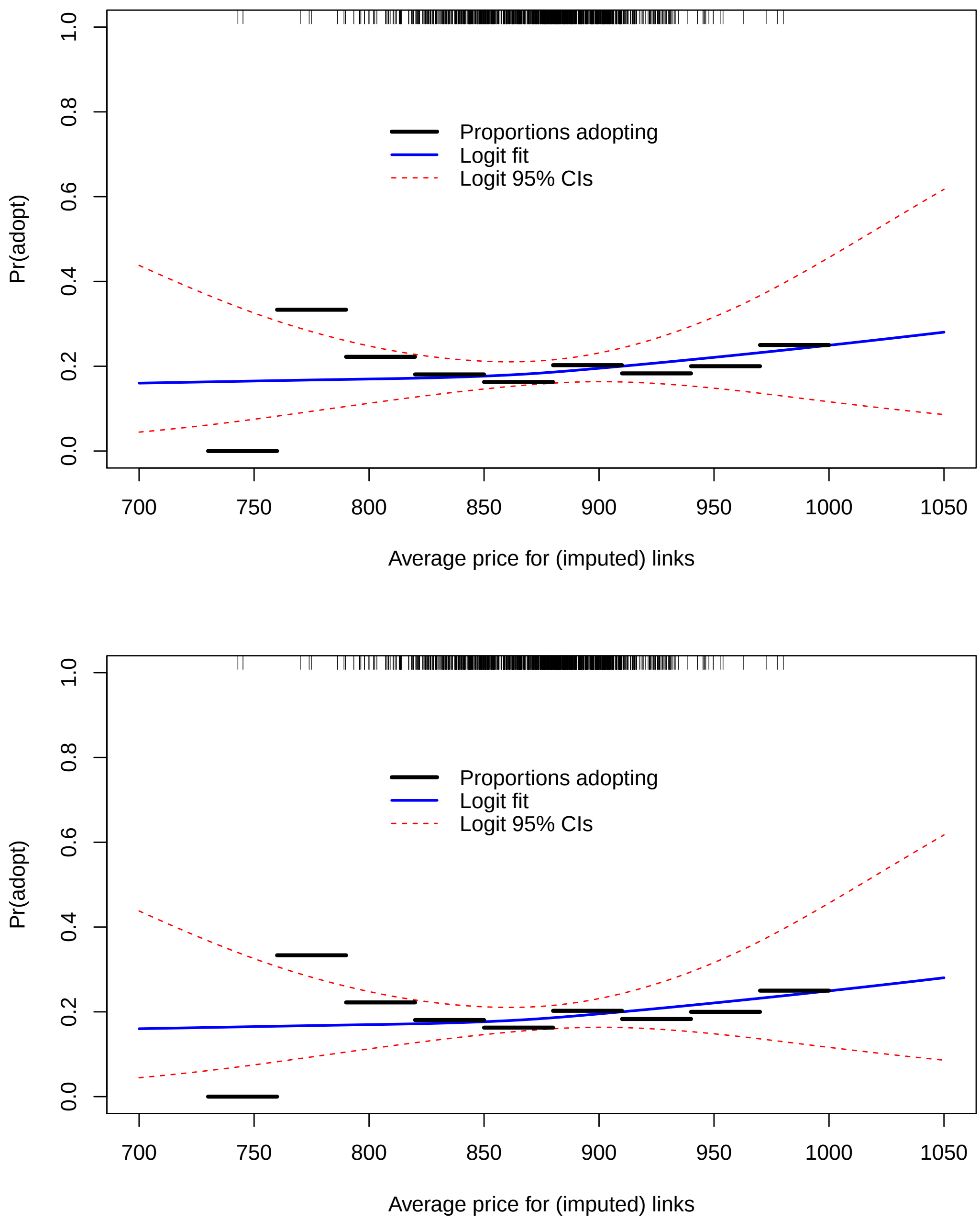
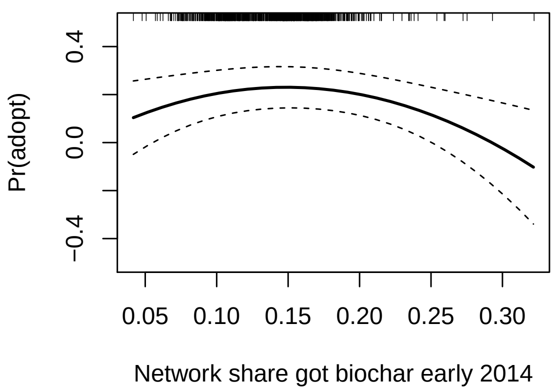


Table C.8: A 2-stage least squares version of model 4.6. First stage models in columns 1-4 – models 1 and 3 are first stages for model 5, while 2 and 4 are first stages for model 6. Standard errors based on White's heteroskedasticity-robust variance-covariance matrix estimate. Plots of the estimated marginal effect of the social network and of price (based on model 6) given below the table. In addition to reported coefficients, all models control for area dummies, gender, having a demonstration plot in late 2013, having heard of ACON, the Rutuba, or Re:char, ever having used charcoal as fertilizer, previous-month income, and RFT repayment discount.

Dependent Variable:	Had biochar early 2014		Share w biochar early 2014		Adopt late 2014	
Intercept	0.08	0.15	0.01**	−0.02*	0.33***	0.04
	(0.13)	(0.19)	(0.00)	(0.01)	(0.09)	(0.13)
Got demo plot	0.08	0.08	0.00	0.00	−0.05	−0.05
	(0.04)	(0.04)	(0.00)	(0.00)	(0.04)	(0.04)
Share net. w/demo	−0.19	−0.33	0.00	0.00	−0.14	0.06
	(0.51)	(0.53)	(0.03)	(0.02)	(0.65)	(0.67)
Late 2014 BC price	−0.01	−0.01	0.00	0.00	−0.09***	−0.09***
	(0.01)	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)
Share with	0.57	−0.53	0.97***	1.64***		
	(0.82)	(3.04)	(0.04)	(0.17)		
Share with cheap offer	2.16	3.87	0.09	0.73**		
	(1.62)	(3.29)	(0.08)	(0.23)		
Risk-free trial	0.65***	0.65***	0.00	0.00		
	(0.07)	(0.07)	(0.00)	(0.00)		
Early 2014 Price	−0.01***	−0.01***	0.00	0.00		
	(0.00)	(0.00)	(0.00)	(0.00)		
Share with cheap offer × Share with	−13.12	−34.49	0.08	−11.14**		
	(11.22)	(50.13)	(0.65)	(3.69)		
Share with cheap offer × (Share with)^2		60.85		45.16**		
		(196.95)		(14.33)		
Early 2014 price x RFT	−0.04***	−0.04***	0.00	0.00		
	(0.01)	(0.01)	(0.00)	(0.00)		
(Share with)^2		4.91		−2.84***		
		(12.97)		(0.74)		
Had biochar early 2014					−0.04	−0.05
					(0.06)	(0.06)
Share with biochar early 2014					−0.24	3.30*
					(0.40)	(1.33)
(Share with biochar early 2014)^2						−11.10**
						(3.99)
R^2	0.72	0.72	0.95	0.96	0.13	0.14
Adj. R^2	0.71	0.71	0.95	0.95	0.11	0.11
Num. obs.	803	803	803	803	803	803

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$



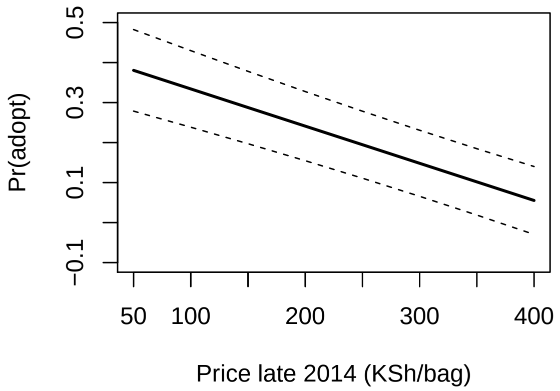


FIGURE C.13: Left: impressions of the difference in yield between portions of farms with biochar, and portions without. Right: distributions of adjusted yield, by stated impression of impact.

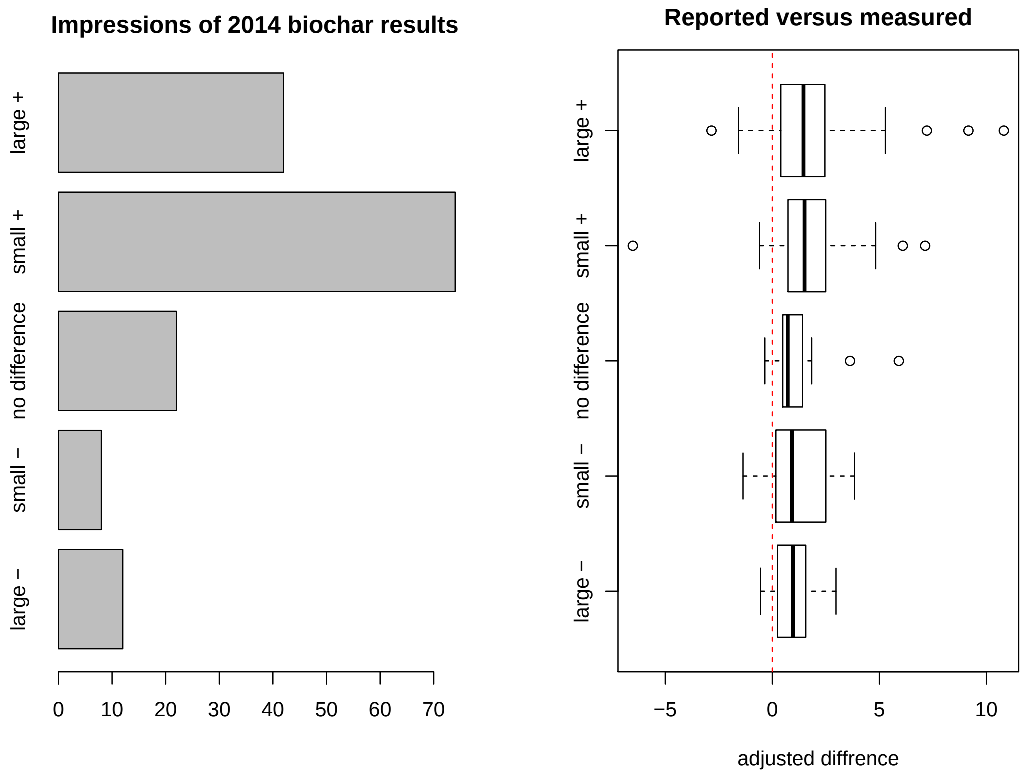


FIGURE C.14: Upper panel: Reported versus measured yields and land under maize for the portion of the sample adopting biochar in early 2014. Lower panel: range of reported maize harvests (reported quantities of maize divided by reported land are under maize) by individual respondents who were surveyed multiple times, for the late-2013 and early 2014 harvest seasons.

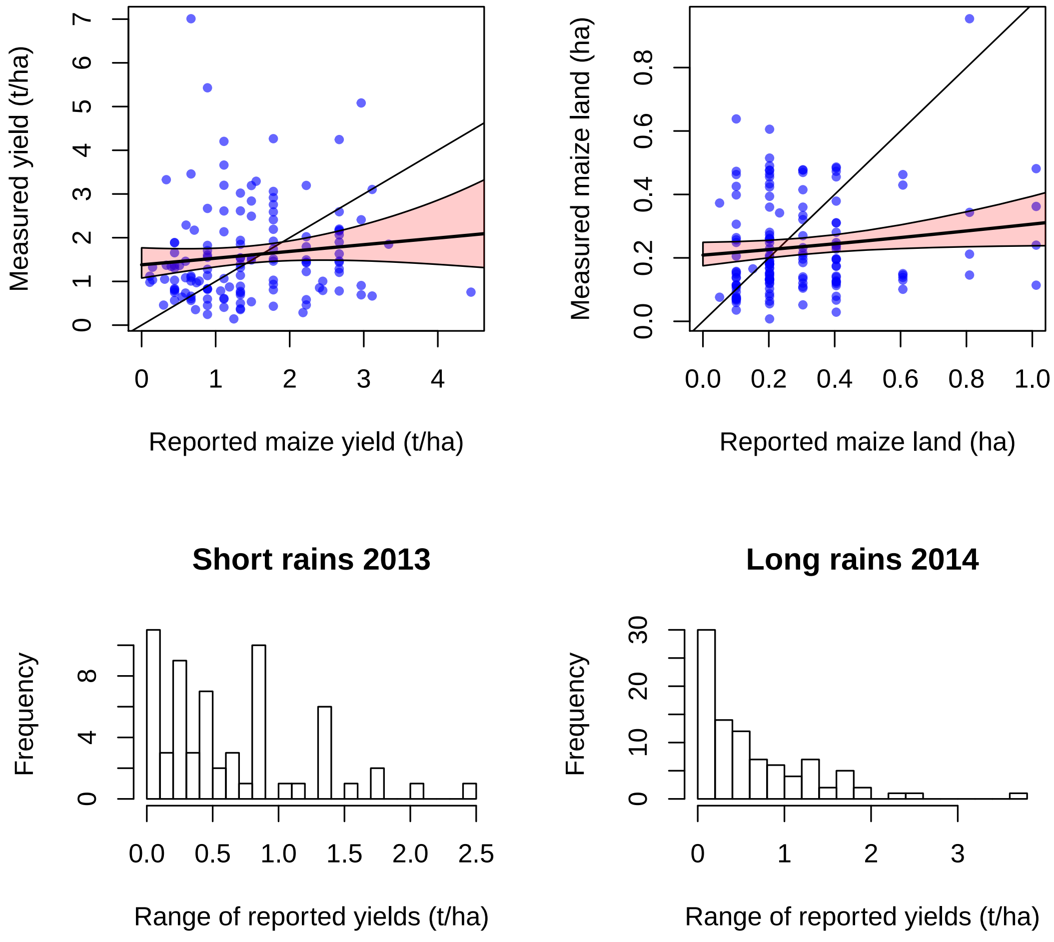


FIGURE C.15: Estimated average response of maize yields to inorganic fertilizer expenditure in the long rainy season, from a fixed-effects regression using data from the long rains of 2010 through the long rains of 2013.

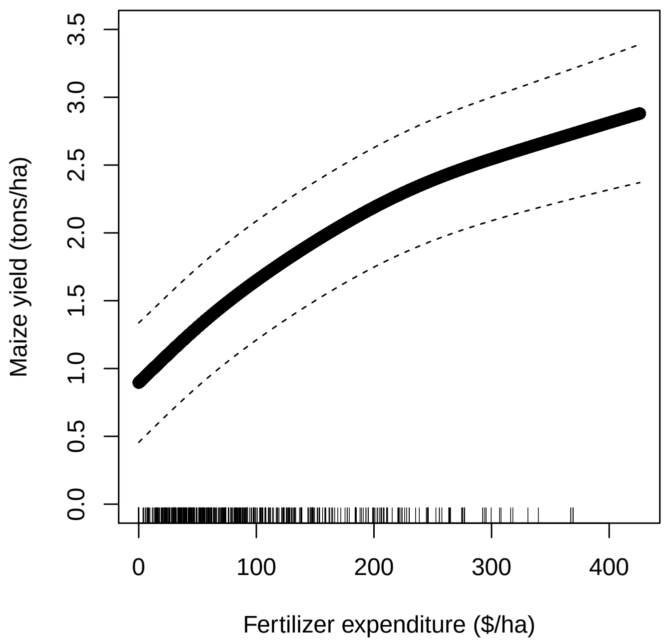


FIGURE C.16: Confidence regions for the combination of subsidies and demonstration plot provision that maximizes adoption at a giver per-farmer program expenditure. Heatmaps give bivariate kernel densities of this combination, based on estimates of propensity to adopt as a function of these treatments (model 4.6), along with modeling assumptions described in section 4.4.5. Calculations based on maximum likelihood estimates of model 4.6 are given by blue dots, and mirror the line shown in figure 4.6b. Confidence regions for these estimates are calculated by taking 500 draws from the posterior of model 4.6; $\boldsymbol{\beta} \sim \mathcal{N}\left(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\beta}}}\right)$, and using those simulates in place of the central estimate in the objective function. 95% confidence regions are drawn to include the highest-density coordinates that contain 95% of the plot's density.

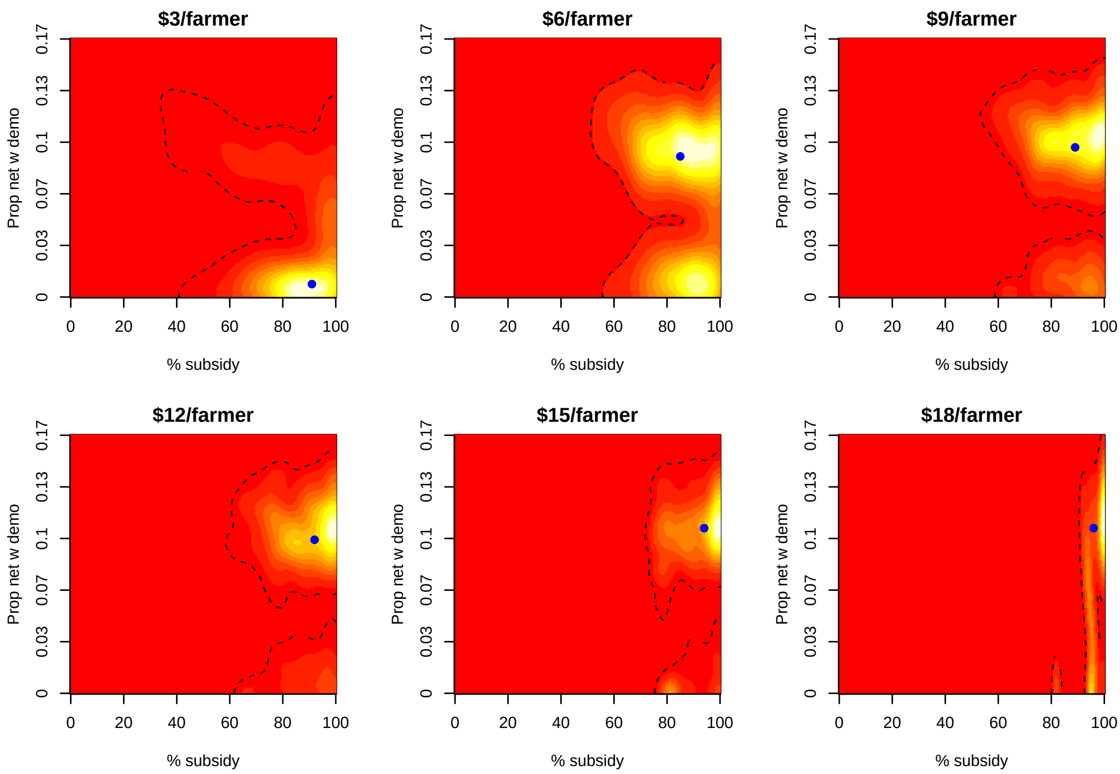


Figure C.17: Confidence regions for point estimates given in figure 4.7, left, in which it is assumed that the median farmer holds fertilizer expenditure constant when adopting biochar. Blue dots give estimates of optimal fertilizer and biochar inducement policy based on maximum likelihood estimates of model 4.6 and figure 4.3. Uncertainty in these estimates is represented by taking 500 draws from the posterior of these models; $\boldsymbol{\beta} \sim \mathcal{N}\left(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\beta}}}\right)$, and using those simulates in place of the central estimate in the objective function. Heatmaps are kernel densities of this output, and 95% confidence regions are drawn to include the highest-density coordinates that contain 95% of the plot's density.

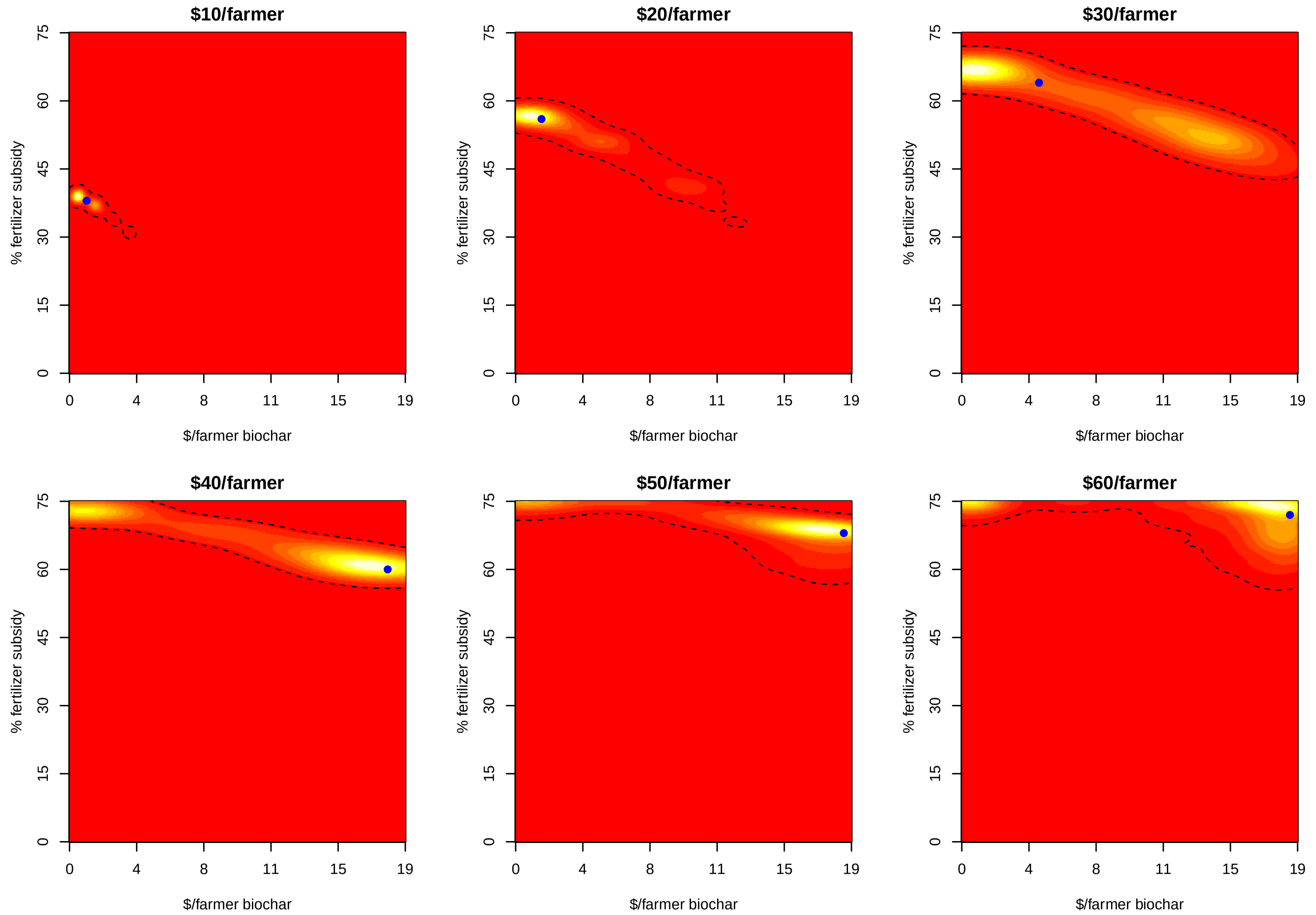


FIGURE C.18: Confidence regions for point estimates given in figure 4.7, right, in which it is assumed that the median farmer holds total input expenditure (biochar plus fertilizer) constant when adopting biochar. Blue dots give estimates of optimal fertilizer and biochar inducement policy based on maximum likelihood estimates of model 4.6 and figure 4.3. Uncertainty in these estimates is represented by taking 500 draws from the posterior of these models; $\boldsymbol{\beta} \sim \mathcal{N}\left(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\beta}}}\right)$, and using those simulates in place of the central estimate in the objective function. Heatmaps are kernel densities of this output, and 95% confidence regions are drawn to include the highest-density coordinates that contain 95% of the plot's density.