Approaches to the study of poverty and environmental impacts of conservation interventions

Bowy den Braber

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

The University of Sheffield

Faculty of Science

Department of Animal and Plant Sciences

September 2019

General summary

Reducing poverty and halting biodiversity loss are two crucial global goals. Protected areas (PAs) are an important example of how goals to reduce poverty and halt biodiversity loss interact. PAs aim to conserve biodiversity, but also have impacts on poverty. In this thesis I focus on the environmental and social impacts of protected areas using a suite of large datasets. Specifically, I focus on how our understanding of PA impacts can be improved by (1) assessing heterogeneity in more detail, (2) comparing impacts relative to impacts of other land uses, and (3) by using better data to study impacts in countries with currently insufficient data. In chapter 2 I assess how Nepali PAs influence poverty, extreme poverty, and inequality. I find that Nepali PAs reduced overall poverty and extreme poverty, and crucially, did not exacerbate inequality. I also find that tourism was a key driver in poverty reductions, but PAs also reduced extreme poverty in areas with few tourists. In chapter 3 I compare PA impacts relative to competing land uses and find that sustainable use PAs, agriculture and mining have led to different outcomes in forest cover and poverty in the Brazilian Amazon. I also show that PAs were effective in reducing deforestation compared to larger-sized landholdings, but not smallholders. I also show evidence that mining sites had more deforestation, but that mining sites also raised local income. In chapter 4 I test whether machine learning methods can be informative to estimate poverty in Tanzania using publicly available satellite imagery. I find that our machine learning methods can be used to estimate household consumption fairly accurately, but cannot be used to measure poverty change or multidimensional poverty. Combined, my findings highlight that PAs can reduce poverty and protect forests although impacts are highly heterogeneous and further scrutiny of PA impacts is needed in more countries. Novel methods using publicly available secondary data show promise to drastically improve the evidence base of PA impacts in data-poor countries where poverty is most prevalent.

Acknowledgements

A huge thanks first of all to my supervisors Karl Evans and Johan Oldekop for their invaluable support and guidance during my PhD. I would also like to thank the wonderful people in the APS department, especially the Conservation Bites group and pub club, for their input, feedback, support, AND Friday fun moonshine times in Interval.

Thanks also go to the Grantham Center for funding my research and broadening the PhD experience. Explaining conservation to physicists and social scientists has proven useful to get your messages across coherently and understandable for all. In that light I also thank the SIID political ecology group for allowing a quantitative person in their mind-boggling discussions.

I also want to thank Rob Hardie, Roxanne Leberger, Prea Singh and Cameron Parker Tas for useful comments on chapters. They have made me realize how blind I can be sometimes.

I had to look Sheffield up on the map, but what a great place it turned out to be, filled with amazing people. In particularly I should mention the people who often joined me on lunches in the deli: Roja, Sheen, Anas, Maha… I will miss our daily lunch rituals and our evening parties. The games! The dancing! The yoga! You guys are crazy cool. I am looking forward to do more fun stuff together in Sheffield and beyond.

Thank you is not enough for my parents who have been there all the way, and who suffered during endless jungle walks because their son wanted to see a boat-billed heron. Sorry dad I did not know I would see (many) more in the years after, but I will never forget the first one…nor your terrified face crawling over a fallen log. I think my mom especially was very happy to see me crunching data instead of wading through crocodile rivers (although I now personally think the latter is easier and safer). I think you have no idea what I really do here in Sheffield but your care and enthusiasm has always been crucial. I am honoured to become the second PhD after Frans. Thank you loads mom, dad, Nelleke, Frans and Jo.

Finally, a huge hug for Bia, who has the magic to make every day a wonderful day. You make me happy, calm, and tired while running after you in countless activities. I often wonder where you find the energy and passion. You have broadened my horizon like no other and life often feels limitless with you. You possess a powerful skill to conquer every obstacle in life, every fallen log, with incredible grace. We have been incredibly lucky with jaguars, tigers and tapirs and I can't wait to find some more animals together! Let's GOOOO!

Author’s declaration

The work presented herein is my own, but owes substantially to the guidance and intellectual contributions of my supervisors Karl L. Evans and Johan A. Oldekop, as well as collaborators for my third chapter Christoph Nolte and Javier Godar, and the useful discussions with Casey Ryan, Gary Watmough and Mauricio Alvarez for my fourth chapter. This thesis has not been submitted to any other reward at this or any other institution.

Chapter 2 has been published as:

Den Braber, B., Evans, K.L., and Oldekop, J.A. (2018) Impacts of protected areas on poverty, extreme poverty and inequality in Nepal. *Conservation Letters*, **11**, e12576.

K.L.E. and J.A.O. are joint last authors.

Chapters 3 and 4 are in preparation and currently unsubmitted. My supervisors provided guidance and support.

Table of contents

Chapter 1:

General introduction

1.1 Global challenges

Eradicating extreme poverty and halting biodiversity loss are two important global challenges to reach the aims of the Sustainable Development Goals (SDGs; Griggs et al. 2013). In this thesis I focus on the poverty and environmental impacts of conservation interventions with a focus on protected areas (PAs). PAs provide a perfect example to study these impacts because they are a key strategy to protect biodiversity but can also have positive and negative socio-economic impacts on local populations (Watson et al. 2014; Brockington & Wilkie 2015). Although PAs are a cornerstone of global biodiversity conservation efforts, much remains unknown about the effectiveness of PAs in halting biodiversity loss, and whether PAs have net positive or negative impacts on local poverty. The lack of evidence in understanding how conservation interventions work has led to numerous calls from scientists to improve the evidence base for more effective policy design and intervention (Fisher et al. 2014; Lu et al. 2015).

1.1.1 Global poverty trends

Before 1750, poverty was largely accepted as a constant feature in society. Poverty reduction was only considered feasible, once Adam Smith's views emerged that economic development was possible (Lipton & Ravallion 1995). This led to a recognition that the state had a role to play in the reduction of, resulting in the poor laws in Britain that were designed to provide safety-nets for the country's poor (Robinson 2002). The earliest estimates of poverty date back to 1820, with more than 80% of the global population living in extreme poverty (earning less than $1 per day; Bourguignon & Morrisson 2002). Yet these measures are crude, because they only estimate poverty by calculating the proportion of GDP that goes to a households. Only from 1980 did the World Bank start data collection to estimate poverty (household consumption expenditure) directly (Grosh & Glewwe 1998). Household consumption expenditure estimates whether a household can secure a basket of goods that is critical for survival, and how much such a basket would cost. Using these data, we now estimate that extreme poverty rates have fallen from 44% of the world's population in 1980 to 10% in 2015 (equivalent to 736 million people earning less than the extreme poverty line of $1.90 per day). The $1.90 poverty line is arbitrary, but poverty rates have decreased regardless of the chosen poverty line (Ravallion 1998). Despite the progress in reducing extreme poverty, an estimated

Met opmerkingen [b1]: Elaboration of examples

3.4 billion people (46% of the world's population) still earn less than $5.50 (the poverty line used in upper-middle-income countries) and struggle to meet basic needs.

Household consumption expenditure remains one of the most used indicators of poverty, but its strictly monetary focus has also been criticized frequently (Alkire & Foster 2011; Brockington et al. 2018). Multiple other dimensions of poverty have been identified, including non-monetary dimensions such as health, education, social inclusion, empowerment and human rights. Several qualitative studies have interviewed the poor to identify their values and identify important dimensions of wellbeing (Narayan 2000).

Met opmerkingen [b2]: Discussion on the various definitions of poverty

Discussion about the definition of poverty have also affected discussions on how poverty should be alleviated or reduced (Agrawal & Redford 2006). From an economic perspective poverty reduction strategies rely on macroeconomic growth, income generation and infrastructure building. However, it has been recognized that removing political, social and cultural obstacles (e.g. addressing neo-colonial power relations) are also important components of poverty reduction strategies.

Much of the criticisms on monetary-based poverty measures draw on the works of Amartya Sen who argued against the use of income-based poverty lines (Anand & Sen 1997; Sen 2000). Firstly, a simple line that divides the world in poor and non-poor obscures important heterogeneity in the intensity of poverty below the poverty line and neglects inequalities between rich and poor. Secondly, counting the poor does not take into account the volatility of poverty (Barrett et al. 2001). Chronically deprived families that are trapped in poverty would require a different approach than families that cycle in and out of poverty. Finally, personal and environmental factors of individuals are not taken into account. In Sen's view income is only a means to an end. Poverty analyses in Sen's view should focus on the ability of people to choose and live the life they value. For example, Sen argued that the right to vote is meaningless if a person has no means of transport to get to the voting booth. This approach is referred to as the capabilities approach and focuses on the actual quality of life that individuals are being able to achieve.

Several attempts have been made to operationalize the capabilities approach into quantifiable indicators. The human development index (HDI) focuses on three capabilities that are regarded as universally valued and weighted equally: long and healthy life (measured by life expectancy at birth), knowledge (measured by expected and mean years of schooling) and decent standard of living (measured by GNI per capita). The HDI has been criticized because the index implies that trade-offs are possible. For example, the HDI implies that an additional year of life has an economic value. The HDI also neglects inequalities because the HDI is developed for the national level. Average income could be high, but taken up only by the elite.

More recently, the multidimensional poverty index (MPI) has been developed that addresses some of these critiques (Alkire & Foster 2011). The MPI is also based on Sen's capability approach and measures poverty in three dimensions: health, education and living standards. However, instead of the HDI, the MPI is calculated at the household level and does not only measure the poverty headcount (whether the household is poor), but also measures poverty intensity (how poor the household is). The more sophisticated calculation of the MPI ensures that trade-offs are not obvious, but the MPI is still criticized because each dimension is weighted equally. Other approaches, such as the Sustainable Livelihood Approach (SLA; Scoones 1998) or the Nested Spheres of Poverty (NESP; Gonner et al. 2007) approach attempt to provide a holistic framework to analyse poverty (and livelihoods in general) by combining information on subjective wellbeing, resources and context.

In general, the broadening of poverty to include multiple dimensions has given rise to a wide range of indicators to determine how poor a person is. These poverty dimensions are not independent, yet the links between them can be complex. For example, marginalization in education could lead to lower incomes and poorer health. Yet higher incomes, especially in rural areas, do not uniformly lead to better drinking water and access to better schools and hospitals. Because of the complex mechanisms that underlie the multiple definitions of poverty, each single poverty indicators remains flawed. Furthermore, several dimensions of poverty are not easily measurable quantitatively, including important dimensions of poverty such as gender inequality, social inclusion and human rights (Narayan 2000). It is therefore important to acknowledge the shortcomings of poverty indicators and to be explicit in what they measure and what they overlook. This is especially important in measuring the success of poverty reduction strategies. Impacts of poverty reduction programmes highly depend on the aspect (incidence, intensity, volatility) and dimension (income, education, health, empowerment) that is taken into consideration (Agrawal 2014).

1.1.2 Global biodiversity trends

In contrast to global poverty estimates, biodiversity trends follow an opposite pattern. Rates of species extinction are 100 to 1,000 times greater than pre-human levels (Pimm et al. 2014). Biodiversity loss is continuing at such alarming rates that some scientists estimate that we are entering the sixth mass extinction if all species currently listed as Critically Endangered go extinct (Barnosky et al. 2011).

The most important drivers of biodiversity loss are land use change, climate change, exploitation (hunting and wildlife trade) and invasive species. Although the importance of each driver varies with taxonomic group and specific location, habitat loss due to land use changes is likely responsible for the largest decline of species. Powers and Jetz (2019) estimate that by 2070 approximately 1,700 species will be imperilled due to land use change. Amphibians are projected to be the most affected taxonomic group as the Red List threat status of 886 amphibian species is projected to be up-listed.

Met opmerkingen [b3]: overall drivers and patterns of biodiversity change

In the tropics forest loss is the main cause of biodiversity loss. Betts et al. (2017) estimate that 121-219 species will become threatened in the next 30 years in Borneo, Central Amazon and the Congo basin. Forest loss also has indirect effects including increasing forest fragmentation and edge effects (Pfeifer et al. 2017), opening up previously inaccessible areas for hunting and wildlife trade, and driving climate change by increasing greenhouse gas emissions (Foley et al. 2005). Forest degradation (negative anthropogenic changes that do not lead to complete forest cover loss) might be less severe than complete deforestation, but is more widespread (Barlow et al. 2016; Edwards et al. 2019). Around 25% of the world's remaining forests are zoned for selective logging, most of which is unsustainable (Edwards et al. 2014). Selective logging typically targets the largest and oldest trees, opening gaps that are often occupied by non-tree species such as climbers. Nevertheless, logged forests are able to provide habitat to a greater proportion of biodiversity than converted areas.

Climate change is causing species extinctions as high-elevation species lose suitable climate conditions on mountains and other species are trapped in forests with unsuitable climatic conditions because of habitat fragmentation and lack of available connections to move between forest patches (Senior et al. 2019). Newbold (2018) estimates that the added effects of climate change and land use changes could lead to a loss of 37.9% of species from vertebrate communities by 2070. Furthermore, by 2070 the effects of climate change could match or exceed the impacts of land use changes.

The effects of overexploitation and invasive species are less well documented. A recent analysis shows that 18% of species are currently being traded, and potentially more species will be traded in the future (Scheffers et al. 2019). Benitez-Lopez et al. (2017) demonstrated that bird and mammal abundances declined due to hunting. Invasive predators have contributed to the majority (58%) of modern bird, mammal, and reptile extinctions (Doherty et al. 2016). The impact of invasive species is especially severe in insular regions, the habitat of a large proportion of highly threatened vertebrate species (Spatz et al. 2017).

In 2010 the Convention for Biological Diversity (CBD) set out twenty “Aichi Biodiversity Targets” to support the overall goal to halt biodiversity loss by 2020, some of which have been adopted by the SDGs. The Aichi targets comprise a wide range of actions that should be taken, including establishing protected areas. Although small but significant progress has been made to achieve some targets, the main goal will not be met (Tittensor et al. 2014).

1.1.3 Linkages between poverty and biodiversity

Although there is increasing recognition that economic growth is not conducive to sustainable development (Stiglitz et al. 2017), at least part of the successes in poverty reductions have been attributed to economic development, especially agricultural development (Ravallion 2001; Deaton 2005). Several Asian countries have shown that reducing poverty can be very successful if agricultural growth is combined with equal distribution of wealth (Chen & Ravallion 2008). However, poverty reduction can occur without economic development (e.g. in Brazil - Ferreira et al. 2010), and agricultural intensification can also have perverse outcomes on well-being (Rasmussen et al. 2018). Nevertheless, agricultural development has been and remains a key policy to promote rural poverty reductions. Yet from a biodiversity perspective, growing demand for agricultural land is also the most important driver of biodiversity loss (Newbold et al. 2015). Most of the recent deforestation has occurred in tropical regions, which is also where most of low and middle-income countries (LMIC) countries are located. Around 150 million hectares of forest have been converted to agriculture between 1980 and 2012, especially in the Amazon and South-East Asia (Hansen et al. 2013). In temperate regions reforestation of previously cleared areas has taken place, causing a rise in forest cover between 1982 and 2016 (Song et al. 2018). Reforestation is considered an important factor in mitigating climate change and restoring biodiversity (Bastin et al. 2019; Busch et al. 2019). The Bonn challenge aims to restore 350 million hectares of forest by 2030. Tracking forest cover estimates and economic development levels have shown a transition starting with high forest cover, deforestation in early stages of economic development followed by reforestation in the latest stages of economic development. This pattern is known as the forest transition curve and suggests that forest cover can regrow after some critical threshold of economic development has been reached (Mather 1992; Meyfroidt & Lambin 2011).

Poverty is also linked to wildlife exploitation. Poor communities greatly depend on forest resources for their livelihoods. For example, bushmeat forms an important component of household income for impoverished communities in the Congo basin (De Merode et al. 2004). However, a study from across Africa found the majority of bushmeat is consumed by wealthier households (Brashares et al. 2011). Bushmeat can also be an important source of nutrients. For example, a study in Madagascar estimated that removing access to bushmeat would triple anemia cases for children living in the poorest households (Golden et al. 2011). These studies show that unsustainable harvest of forest resources and the subsequent depletion of wildlife could also increase poverty.

A closer look at the relationship between development and nature could give insights into how sustainable development could be achieved. The consequences of poverty reduction could have negative environmental consequences, but in other cases poverty reductions could synergize with positive biodiversity outcomes. Nevertheless,, sustainable development would likely require a radical reconfiguration of how humans interact with the natural world (Foley et al. 2005; Cardinale et al. 2012). The enormity of this task raises questions about the importance of biodiversity and whether it is worth the effort. Biodiversity is important for humans because it has intrinsic and extrinsic value. Intrinsic value is based on the belief that all life on earth deserves the right to exist. Extrinsic value relates to the benefits that nature provides for humans such as ecosystem services (Costanza et al. 1997; Balmford et al. 2002). A disproportionate part of poor people lives in natural systems and directly depends on the diversity of ecosystem services that nature provides (Jagger et al. 2014). Ecosystems that harbour maximum biological diversity are more resilient to human or natural shocks such as climate change (Isbell et al. 2015). At the local scale poor people will be hit the hardest when nature's resilience decreases. On the global scale exceeding planetary boundaries will lead to severe negative effects for all people (Rockström et al. 2009).

Met opmerkingen [b4]: Expansion on the links between biodiversity and poverty

1.2 Protected areas as coupled human and natural systems

Disentangling the complex interactions in coupled human and natural systems (CHANS) is therefore crucial and requires a holistic approach, yet social and environmental studies have developed mostly independently (Liu et al. 2007; Ostrom 2009). By integrating the environmental and social sciences we have reached a better understanding of the complexity of these systems and how they interact. Scientists of CHANS have identified causal relationships in a wide range of human-nature interactions including deforestation, fishing, desertification, species extinctions, and collapse of ecosystems and civilizations (Ferraro et al. 2018). They have shown a myriad of synergies and trade-offs between poverty reduction and biodiversity conservation efforts. These have sparked debates about the roles of poverty reduction in biodiversity conservation efforts and vice versa (Soule 2013; Marvier 2014; Roe 2019).

Protected areas (PAs), the main focus in this thesis, are an example of a CHANS and play a key role in these discussions. Setting aside land for protection has developed as one of the key strategies of contemporary biodiversity conservation. The concept of PAs has a long history in many cultures, but the first official PA only dates from 1872 with the establishment of Yellowstone NP. As environmental awareness increased governments around the world decided to protect more land (Watson et al. 2014). Currently around 16% of the terrestrial surface falls under protection, close to the Aichi target of 17% that the CBD set to reach by 2020. Although the global PA network is expanding, reports of downgrading, downsizing and degazettement of PAs have increased too, highlighting that the PA network is more dynamic than previously assumed (Golden Kroner et al. 2019). For example, the Arabian Oryx sanctuary in Oman currently only retains 10% of its original size for hydrocarbon purposes. The Yosemite National Park in the USA has experienced multiple changes, including downsizes (for forestry and mining) and downgradings (to allow construction of roads, facilities and dams). Some changes have been reversed later, or were off-setted by protecting other tracts of land (Qin et al. 2019).

Met opmerkingen [b5]: Example to clarify point made

The primary goal PAs is to protect the biological diversity within its borders. PAs can achieve this by mitigating habitat loss, minimizing anthropogenic disturbances and avoiding species extinctions and population declines. PAs also have socio-economic impacts by restricting access and use to resources within. It is estimated that around 70% of the PAs has people living within PA boundaries (Bruner et al. 2001). The world conservation union IUCN classifies PAs into six different categories according to the extent to which biodiversity goals are prioritized over impacts on people (Dudley 2008). Categories I-IV are regularly referred to as Strict PAs because they do not allow human use, while categories V and VI are sustainable use (SU) PAs and explicitly list community use as one of their major objectives. Initially PAs excluded human interference based on Hardin's influential theory "the tragedy of the commons" that hypothesised that people (commoners) were unable to limit their resource use, and should be separated from nature in order to protect it. This paradigm has also been referred to as fortress conservation and was later critiqued as the recognition grew that biodiversity conservation should not be separated from the involvement of local communities (Naughton-Treves et al. 2005; Ostrom & Nagendra 2006; Berkes 2007).

PAs show varied success in achieving their conservation and socio-economic goals. The expansion of the PA network has been celebrated as a key conservation success, but concerns have been raised about the effectiveness of PAs in protecting habitat and species. PA management is plagued with funding shortfalls (Waldron et al. 2017), and PAs are chronically understaffed (Coad et al. 2019). Critics argue that some PAs only exist on paper and are referred to as 'paper parks' (Oates 1999). A PA management assessment of 2167 PAs shows that only a quarter of terrestrial PAs have adequate staffing and budget (Coad et al. 2019). Insufficient resources impact law enforcement, boundary demarcation, natural resource management and could also lead to reduced communication with stakeholders, potentially creating conflicts (Watson et al. 2014). Corruption and bureaucracy are also important factors that influence PA performance and effectiveness. Upgrading the management of current PAs could result in significantly improved conservation outcomes (Pringle 2017; Kuempel et al. 2018). For example, the Parque Nacional de Gorongosa in Mozambique suffered great ecological damage during civil wars (Pringle 2017). More than 90% of the park's large mammals was extinguished in 1992, but wildlife recovered to almost 80% of pre-conflict levels following an upgrade of PA management.

Met opmerkingen [b6]: Example to clarify point made

The success of PAs is not only influence by its management but also by its location. Many PAs protect lands with low opportunity costs, such as areas with higher elevations, steeper slopes and remote from cities (Joppa & Pfaff 2009). Consequently, many PAs only passively protect forests through the absence of human pressure, and thus play a limited role in conservation. PAs are often established in an opportunistic, *ad ho*c manner that generally does not take into account species representation across all biomes. Only 15% of the threatened vertebrates are effectively represented in PAs (adequate overlapping of species range with a PA - Rodrigues et al. 2004; Venter et al. 2014). Policies to expand PAs based on area alone can thus result in underachieving biodiversity outcomes. Area-based targets of PAs are compared to Goodhart's

Law which warns that an indicator can transition into a de facto policy target, thereby neglecting the overarching goal (Barnes et al. 2018).

1.2.1 Impacts of protected areas on biodiversity

Despite concerns regarding PA management and location, studies have shown that PAs have a net positive impact on protecting habitats. Remote sensing to track forests from space has shown that PAs on average reduce deforestation and carbon loss. Examples from countries across all continents show that tropical deforestation rates within PAs are routinely lower in PAs than in comparable sites outside PAs (e.g. Andam et al. 2008; Canavire-Bacarreza & Hanauer 2013; Sims & Alix-Garcia 2017), although Brun et al. (2015) found no impact of PAs in Indonesia. Effect sizes of PA impacts vary per country, per protected area and can even differ within PAs. For example Shah & Baylis (2015) found that the Sebangau National Park in Kalimantan, Indonesia, performed significantly better in reducing deforestation than the Kerinci Seblat National Park in Sumatra. Within the Kerinci Seblat National Park, reducing deforestation was more successful in the South than in the North part of the park. On average, the effect size were modest. For example, Andam et al. (2008) found that PAs avoided 10% of deforestation in Costa Rica. It is possible that PAs also have positive ("blocking") or negative ("leakage") impacts outside the boundaries, but the evidence remains inconclusive (Herrera et al. 2019). Fewer studies show evidence on PA success in other biomes, although a study in the Brazilian Cerrado found that PAs were successful in reducing deforestation (Carranza et al. 2014). The effectiveness of PAs is also influenced by their management goals, but strictly protected areas are not always more effective in reducing deforestation than SU PAs (Ferraro et al. 2013). Despite the relative success of PAs in slowing down deforestation in comparison to unprotected land, 3% of the global protected forest was lost from 2000 to 2012 and 10% of the total forest loss occurred within PA boundaries (Heino et al. 2015). Even in some renowned PAs that have been designated as UNESCO world heritage sites deforestation has continued. For example, the Río Plátano Biosphere Reserve in Honduras lost 5% forest cover between 2000 and 2012 (Allan et al. 2017). In some PAs deforestation rates have been especially alarming. A study from Kalimantan, Indonesia, has shown that protected forests have declined by more than 50% (Curran et al. 2004).

Met opmerkingen [b7]: Example to clarify point made

Met opmerkingen [b8]: Example to clarify point made

Few studies have focused on the impacts of PAs on smaller scale anthropogenic disturbances such as logging and fire. Anthropogenic impacts are often not measured but are severely underestimated (Barlow et al. 2016). Traditional remote sensing techniques are unable to detect selective logging because the pixel size is too big to capture small-scale habitat changes. Using techniques to analyse forest degradation at subpixel levels one study has shown that PAs in the Brazilian Amazon show limited signs of selective logging (Asner et al. 2005). Evidence from other regions is more limited. Fire is an important indicator of land clearing of forests in Asia and South-America, but fires form an integral part of management of grassland ecosystems. A global study compared fires inside PAs vs outside PAs and found fewer fire occurrences in PAs (Nelson & Chomitz 2011). SU PAs had fewer fires than strict PAs.

Evidence of the impacts of PAs on conserving species is still weak. Populations of many species are declining in PAs (Craigie et al. 2010), but many studies of PA impacts on species populations and assemblages lack an appropriate counterfactual to compare PA impacts against. Identifying a counterfactual can be complicated when species mostly occur within PAs or if species migrate beyond PA boundaries. Furthermore, studies on PA impacts on species are often biased towards charismatic species, such as lions and elephants, and well-known PAs with adequate resources (Geldmann et al. 2013). Two meta-analyses report higher species abundance inside PAs than outside, but sample sizes (42 and 86 studies) were limited (Geldmann et al. 2013; Coetzee et al. 2014). The most comprehensive study compared species richness and abundance of 13,669 vertebrate species inside 359 PAs against a counterfactual without protection (Gray et al. 2016). They found higher species richness and abundance inside PAs than outside. However, wide confidence intervals highlighted that evidence of PA impacts on species should be improved. They also found that differences in species were mostly due to differences in habitat instead of wildlife exploitation (e.g. hunting and trapping). The evidence on impacts of PAs on species through limiting exploitation of wildlife is limited. Hunting pressure is generally lower in PAs than outside PAs, but overhunting in PAs is common in tropical PAs (Laurance et al. 2012; Benítez-López et al. 2017). The effects of PAs on mitigating exploitation for wildlife trade are poorly known although studies suggest that PAs might fail to prevent the extinction of commercially valuable species (Symes et al. 2018).

1.2.2 Socio-economic impacts of protected areas

The rules and regulations that PAs impose also have positive and negative socio-economic impacts for people. Socio-economic outcomes can also influence biodiversity outcomes. For example, local poverty is an important factor in explaining elephant poaching (Hauenstein et al. 2019). Oldekop et al. (2016) found that PAs with positive socio-economic outcomes were also more likely to report positive biodiversity conservation outcomes.

PAs can provide positive socio-economic outcomes by safeguarding vital ecosystem services, including water provision, food security and carbon storage. For example, cities often rely on PAs as a source of drinking water (Watson et al. 2014). Globally, PAs are also play an important role in climate change mitigation efforts, such as REDD+ (reducing emissions from deforestation and forest degradation, plus the conservation, sustainable management and enhancement of forest carbon stocks) schemes (Scharlemann et al. 2010).

PAs can also positively aid communities by creating employment opportunities, especially through ecotourism. Ecotourism has been rising in recent years and some countries are now highly dependent on ecotourism as driver of their economy (Coria & Calfucura 2012). Birdwatching trips in Colombia could generate an annual profit of $9 million and create 7516 new jobs (Maldonado et al. 2018). Presence of ecotourism is also linked to higher wages. For example, in Costa Rica Robalino et al. (2015) found higher wages of people living closer to PA tourist entrances. A case study in Cuyabeno National Park, Ecuador, showed that ecotourism has provided substantial additional income, even to families that are not directly involved in tourism activities (Wunder 2000). Studies from other countries confirm that ecotourism can lead to improved income and employment for local communities (Walpole & Leader-Williams 2001), but other studies warn that ecotourism can cause inequalities between communities, pressure on local resources, and even conflicts between PAs and communities (West et al. 2006). For example, in Qinling province in China giant panda reserves reduced poverty, but also increased inequality (Ma et al. 2019).Studies from other countries confirm that ecotourism can lead to improved income and employment for local communities (Walpole & Leader-Williams 2001), but other studies warn that ecotourism can cause inequalities between communities, pressure on local resources, and even conflicts between PAs and communities (West et al. 2006).

Met opmerkingen [b9]: Elaboration including examples to clarify the points made

Of the negative impacts displacement of local people from PAs is certainly one of the most controversial (Cernea & Schmidt-Soltau 2006; Agrawal & Redford 2009). A report on displacement of people in 36 Central African PAs shows that in 26 of them PA creation had resulted in partial or complete displacement of local people (Pyhälä et al. 2016). In no case compensation was provided. The magnitude of displacement is poorly known, but could have affected millions of people (Agrawal & Redford 2009). Yet, much of the evidence is anecdotal and is biased towards certain regions. It is unclear how many evictions will take place to make room for future PAs. A study from India suggests that four million people will be evicted from PAs in the future. Managing forthcoming displacement by the expansion of the PA network is one the key social challenges (Büscher et al. 2017). One study from India shows how local communities were successfully resettled to make room for the Bhadra Wildlife Reserve (Karanth 2007).

Damning reports about the abuse of local people by park rangers in Central Africa have caused further concern about the social impacts of PAs (Warren & Baker 2019). Some have argued that the response of the international NGO community has been muted, which can have dire consequences for international NGOs. The government of Germany has stopped funds for WWF following a flurry of human rights violations by park rangers in Central Africa (Engert et al. 2019).

Such troublesome reports have received significant attention, but the economic displacement that arises through restrictions in use and access of resources is arguably more common (Cernea & Schmidt-Soltau 2006). These restrictions disproportionally affect poor segments of the population who rely on natural resources for their livelihoods (Holmes 2007). Examples from community-based conservation projects have shown how PAs can sustain livelihoods and safeguard ecosystem services , but stories of failure are common too (Weber et al. 2011). There is wide recognition that dialogue between PA managers and stakeholders is crucial, but consultation and participation of communities to guide management is often lacking. Communities were involved in management decisions in only 8 of the 34 surveyed PAs in Central Africa (Pyhälä et al. 2016). In only one PA the participation resulted in concordant management and zoning decisions that benefit community interests. However, in Nepal communities are more actively involved in PA management decisions and 30-50% of PA revenues are distributed to local communities (Heinen & Shrestha 2006).

Positive and negative PA socio-economic impacts can have an impact on poverty of local communities (Adams et al. 2004; Brockington & Wilkie 2015b). Poverty is a complex concept that encompasses multiple dimensions such as income, education, health and household assets (Anand & Sen 1997). Poverty can be measured by using proxies of each dimension or by combining multiple dimensions of poverty into one index, for example the multidimensional poverty index (MPI; Alkire & Foster 2011). Currently there is no consensus on which measures of poverty are best, and studies to measure the impact of PAs on poverty have used multiple different measures of poverty. Nevertheless, a growing literature of impact evaluations of PAs on poverty does not find evidence of poverty exacerbation by PAs, although PAs are often located in areas with high poverty levels (Fisher & Christopher 2007). Instead, several studies show that PA networks in several countries (e.g. Costa Rica, Thailand and Bolivia) have managed to reduce poverty compared to comparable non-protected areas using a wide range of poverty metrics. A global assessment of the impact of PAs did not find negative effects of PAs on human health and living standards (Naidoo et al. 2019). They found that PAs with documented tourism were associated with better children's health and more assets. These results suggest that benefits stemming from ecotourism are beneficial for households and additional income can be spent to buy assets and to pay for food, medicine or medical clinic visits that improve children's health. Ferraro et al. (2015) found that ecotourism was able to explain two-thirds of the reduction in poverty in Costa Rica PAs. There is less evidence whether strict or SU PAs are better at reducing poverty. Globally SU PAs performed better in improving health and assets (Naidoo et al. 2019), but studies within countries found mixed results (e.g. Miranda et al. 2014).

1.2.3 Approaches to measure impacts of protected areas

Multiple qualitative and quantitative studies to assess PA impacts have been implemented across the world (Pullin et al. 2013). However, many study designs faced severe limitations. Past studies generally did not record pre-intervention levels, did not include a control group to compare the impact of the intervention against, and did not randomize allocation of the intervention (Baylis et al. 2015). Instead, most studies have only considered a single site at a single point in time. Such studies have put emphasis on the contextual and place-based impacts, but these impacts might not be representative of impacts in other conditions. Therefore their findings are mostly unsuitable for generalization and are unable to serve as basis for incorporation into broad policy interventions (Agrawal & Redford 2006; Miteva et al. 2012).

Indeed, study designs such as before-after-control-impact (BACI) or randomized control trials (RCTs) that incorporate pre-intervention sampling, counterfactuals, and/or random allocation result in more accurate effect estimations (Christie et al. 2019). However, such study designs can be expensive, infeasible, or even unsuitable for many research questions in conservation and development (Deaton 2009).

Recently, conservation and development scientist have applied a quasi-experimental econometric approach that mimics these study designs to evaluate impacts. By controlling for the influence of potential confounding variables these studies are able to identify treatment and counterfactual control groups without the intervention. This technique is referred to as statistical matching, and allows to empirically quantify the impact of an intervention in comparison to a counterfactual without the intervention (Stuart 2010). Following this methodology studies have examined the environmental and socio-economic impacts of a multitude of policies and interventions in CHANS, including PAs. PA impact evaluations generally combine big datasets to isolate impacts of PAs from land use processes that would have happened anyways as determined by factors that influence the value of land (Meyfroidt et al. 2018), including biophysical characteristics (Ricardo 1891) and accessibility/distance to markets (Von Thünen 1966). More recently, PA impact evaluations have also shown through which mechanisms (e.g. tourism) PAs affect environmental and socio-economic outcomes of PAs (Ferraro & Hanauer 2015), and under which conditions impacts are moderated by exogenous variables (Hanauer & Canavire-Bacarreza 2015; Ferraro et al. 2018).

1.2.4 Data gaps in measuring impacts of protected areas

The large increase of data has greatly advanced research on CHANS. An important example is the wall-to-wall forest map that records global forest cover changes from 2000 (Hansen et al. 2013). Such forest maps allow to evaluate the impacts of entire PA networks. In contrast to these environmental data, data gaps on poverty metrics are still widespread. A few global initiatives collect data on socio-economic indicators in LMIC counties. The World Bank LSMS studies collect data from many countries to estimate global poverty rates based on household consumption expenditure (Grosh & Glewwe 1998). In addition, the demographic household study (DHS) collects data on other dimensions of poverty that are important to people's capabilities, including education, health and living standards. The Oxford Poverty and Human Development Initiative (OPHI) collates these data to estimate the multidimensional poverty indices for countries (Alkire & Foster 2011). The DHS data has been used to measure global PA impacts on health and assets (Naidoo et al. 2019). However, the sampling coverage of both global initiatives is limited to only 1-5% of the entire populations in most countries. In addition, geolocations of sampled households are made fuzzy (up to 10 km) to ensure privacy, and sampled households might not to be representative of the local communities at each location. Therefore, most PA impact evaluations have resorted to census data that are generally recorded in cycles of ten years. Census data cover a much larger percentage of the population (typically around 10%) and are often designed to be representative of local communities and administrative areas. However, censuses are costly endeavours that are often not implemented in LMIC countries, or have been skipped for some years. Therefore, country-wide socio-economic impacts of PAs have mostly been studied in upper middle-income countries (e.g. Costa Rica). As far as I know socio-economic impacts of PAs have only been studied in one LMIC country (Bolivia; Hanauer & Canavire-Bacarreza 2015), and only one country listed as least developed (Nepal; den Braber et al. 2018). No country-wide assessment of PAs has taken place in sub-Saharan Africa, where most of the negative impacts of PAs are reported. Hence, there is a clear need for better data to evaluate socio-economic impacts of PAs and other policies in CHANS.

An alternative to census data is to use remote sensed satellite data that can be used as proxies to measure socio-economic indicators. For example, luminosity (nightlight) measured by satellites can be a proxy for poverty (Noor et al. 2008; Chen & Nordhaus 2011). However, in areas with high poverty, luminosity levels are generally very low, making it difficult to distinguish unpopulated areas from areas with high poverty levels. Machine learning methods have shown promise to derive information from remotely sensed imagery that can be applied to predict poverty (Watmough et al. 2016, 2019). One approach used by Jean et al. (2016) applies convolutional neural networks (CNNs) and imagery from Google Street Maps to detect poverty in four African countries at higher accuracy than using luminosity. CNNs are a type of deep neural networks that are developed specifically for imagery (Lecun et al. 1998; Bengio 2009). In contrast to other machine learning methods they are able to detect spatial features from images. Jean et al. (2016) show that CNNs are able to detect landscape features from imagery that correlate to economic activity and are useful for detecting poverty.

Advancing socio-economic data would allow to estimate the socio-economic impacts of PAs and other policies in countries where data is currently not available, but where poverty levels are most dire. Studying impacts of policies on habitat and poverty simultaneously would advance our understanding of interactions in CHANS. These insights could be used to develop strategies that combine poverty reduction and halt biodiversity loss to achieve the SDGs.

1.3 Thesis aims and objectives

In this thesis I focus on approaches in the study of environmental and social impacts of protected areas. Specifically, I focus on how our understanding of PA impacts can be improved by (1) assessing heterogeneity in more detail, (2) comparing impacts relative to impacts of other land uses, and (3) by filling data gaps to study impacts in countries with currently insufficient socio-economic data.

In chapter 2, I assess how Nepalese PAs influence poverty, extreme poverty, and inequality. Despite progress in assessing PA impacts on poverty, studies have mostly focused on entire communities. Yet benefits of PAs might be captured by the elite, potentially exacerbating inequality (West et al. 2006). Previous studies have suggested that tourism could be an important mechanism through which PAs impact local poverty by providing local income and employment (Ferraro & Hanauer 2014). Studies have used binary assessments of tourism (presence or absence of tourism), but there is a need to move beyond binary impacts to gain a better understanding of how tourism affects local poverty. Finally, PA impacts could be moderated by duration of protection, the amount of protected land and elevation, an important factor in determining agricultural suitability (Joppa & Pfaff 2009).

This chapter sets out to assess how PAs in Nepal influence different metrics of poverty, using a multidimensional poverty index and a quasi-experimental design that controls for potential confounding factors. I specifically assess the role of tourism by assessing PA impacts according to the number of tourists they receive, and by assessing the role of tourism infrastructures like trekking routes.

Chapter 3 addresses the effects of PAs in the Brazilian on deforestation and poverty. Specifically, I assess the impact of PAs in comparison to two important competitors of land use: agriculture and mining. Policy makers have to continually make trade-offs between environmental and social goals. Quantifying the trade-offs and synergies between different land use options could be crucial (Lambin et al. 2014), yet impact evaluations rarely go beyond binary assessments of land uses (e.g. protection vs not protection). Due to looming land scarcity (Lambin & Meyfroidt 2011), competition for land is expected to increase, and will also affect protection as already shown by the downgrading, downsizing and degazettement of PAs worldwide (Golden Kroner et al. 2019), and also in the Amazon (Pack et al. 2016).

In this chapter I analyse whether sustainable use PAs, agriculture and mining lead to different outcomes in forest cover and poverty in the Brazilian Amazon. I chose SU PAs because they allow some access and use within the PA and are therefore more comparable to extractive land uses. I use a quasi-experimental study design that controls for confounders. I split agriculture according to different landholding sizes and assess spread of mining impacts at different buffer sizes.

Chapter 4 tests the ability of novel machine learning methods to measure poverty in Tanzania. A better understanding of how interventions impact poverty requires quantitative assessments at local scales and different time points (Lu et al. 2015). Because household surveys are costly endeavours that are unfeasible in many countries, using remotely sensed data from satellites could be a promising alternative to close the data gap (Watmough et al. 2019). Recently developed machine learning algorithms (convolutional neural networks (CNNs)) show that they can estimate poverty better than traditional proxies of poverty such as nightlights (Jean et al. 2016). However, these studies used very high resolution imagery that is not publicly available at large spatial scales.

In this chapter I test whether CNNs could be informative to estimate poverty using publicly available satellite imagery. I compare predictive power to estimate two different metrics of poverty (household consumption expenditure and a multidimensional poverty index) and assess if CNNs can also accurately estimate consumption change between two different time points.

Finally, in chapter 5 I synthesise my findings, and discuss the value for policy makers and PA managers of assessing PA impacts. I also show how future impact evaluations could provide more insight into PA impacts and I give several recommendations for future research.

1.4 References

Adams, W.M., Aveling, R., Brockington, D., Dickson, B., Elliott, J., Hutton, J., Roe, D., Vira, B. & Wolmer, W. (2004). Biodiversity conservation and the eradication of poverty. *Science (80-.).*, 306, 1146–1149.

Agrawal, A. & Redford, K. (2006). Poverty, development, and biodiversity conservation: Shooting in the dark? *Ann Arbor MI*, 48109, 647–5948.

Agrawal, A. & Redford, K. (2009). Conservation and displacement: An overview. *Conserv. Soc.*

Alkire, S. & Foster, J. (2011). Counting and multidimensional poverty measurement. *J. Public Econ.*, 95, 476–487.

Allan, J.R., Venter, O., Maxwell, S., Bertzky, B., Jones, K., Shi, Y. & Watson, J.E.M. (2017). Recent increases in human pressure and forest loss threaten many Natural World Heritage Sites. *Biol. Conserv.*, 206, 47–55.

Anand, S. & Sen, A. (1997). Concepts of Human Development and Poverty: A Multidimensional Perspective. *United Nations De-velopment Program. Poverty Hum. Dev. Hum. Dev. Pap.*, 1–20.

Andam, K.S., Ferraro, P.J., Pfaff, A., Sanchez-Azofeifa, G.A. & Robalino, J.A. (2008). Measuring the effectiveness of protected area networks in reducing deforestation. *Proc. Natl. Acad. Sci. U. S. A.*, 105, 16089–16094.

Asner, G.P., Knapp, D.E., Broadbent, E.N., Oliveira, P.J.C., Keller, M. & Silva, J.N. (2005). Selective logging in the Brazilian Amazon. *Science (80-.).*, 310, 480–482.

Balmford, A., Bruner, A., Cooper, P., Costanza, R., Farber, S., Green, R.E., Jenkins, M., Jefferiss, P., Jessamy, V., Madden, J., Munro, K., Myers, N., Naeem, S., Paavola, J., Rayment, M., Rosendo, S., Roughgarden, J., Trumper, K. & Turner, R.K. (2002). Ecology: Economic reasons for conserving wild nature. *Science (80-.).*

Barlow, J., Lennox, G.D., Ferreira, J., Berenguer, E., Lees, A.C., Nally, R. Mac, Thomson, J.R., Ferraz, S.F.D.B., Louzada, J., Oliveira, V.H.F., Parry, L., Ribeiro De Castro Solar, R., Vieira, I.C.G., Aragaõ, L.E.O.C., Begotti, R.A., Braga, R.F., Cardoso, T.M., Jr, R.C.D.O., Souza, C.M., Moura, N.G., Nunes, S.S., Siqueira, J.V., Pardini, R., Silveira, J.M., Vaz-De-Mello, F.Z., Veiga, R.C.S., Venturieri, A. & Gardner, T.A. (2016).

Anthropogenic disturbance in tropical forests can double biodiversity loss from deforestation. *Nature*, 535, 144–147.

Barnes, M.D., Glew, L., Wyborn, C. & Craigie, I.D. (2018). Prevent perverse outcomes from global protected area policy. *Nat. Ecol. Evol.*

Barnosky, A.D., Matzke, N., Tomiya, S., Wogan, G.O.U., Swartz, B., Quental, T.B., Marshall, C., McGuire, J.L., Lindsey, E.L., Maguire, K.C., Mersey, B. & Ferrer, E.A. (2011). Has the Earth's sixth mass extinction already arrived? *Nature*.

Bastin, J.F., Finegold, Y., Garcia, C., Mollicone, D., Rezende, M., Routh, D., Zohner, C.M. & Crowther, T.W. (2019). The global tree restoration potential. *Science (80-.).*, 364, 76–79.

Baylis, K., Honey-Rosés, J., Börner, J., Corbera, E., Ezzine-de-Blas, D., Ferraro, P.J., Lapeyre, R., Persson, U.M., Pfaff, A. & Wunder, S. (2015). Mainstreaming impact evaluation in nature conservation. *Conserv. Lett.*

Bengio, Y. (2009). Learning Deep Architectures for AI. *Found. Trends® Mach. Learn.*, 2, 1–127.

Benítez-López, A., Alkemade, R., Schipper, A.M., Ingram, D.J., Verweij, P.A., Eikelboom, J.A.J. & Huijbregts, M.A.J. (2017). The impact of hunting on tropical mammal and bird populations. *Science (80-.).*, 356, 180–183.

Berkes, F. (2007). Community-based conservation in a globalized world. *Proc. Natl. Acad. Sci. U. S. A.*, 104, 15188–15193.

Betts, M.G., Wolf, C., Ripple, W.J., Phalan, B., Millers, K.A., Duarte, A., Butchart, S.H.M. & Levi, T. (2017). Global forest loss disproportionately erodes biodiversity in intact landscapes. Nature, 547, 441–444.

Bourguignon, F. & Morrisson, C. (2002). Inequality among world citizens: 1820-1992. *Am. Econ. Rev.*, 92, 727–744.

den Braber, B., Evans, K.L. & Oldekop, J.A. (2018). Impact of protected areas on poverty, extreme poverty, and inequality in Nepal. *Conserv. Lett.*, e12576.

Brashares, J.S., Golden, C.D., Weinbaum, K.Z., Barrett, C.B. & Okello, G. V. (2011). Economic and geographic drivers of wildlife consumption in rural Africa. Proc. Natl.

Acad. Sci. U. S. A., 108, 13931–13936.

Brockington, D. & Wilkie, D. (2015a). Protected areas and poverty. *Philos. Trans. R. Soc. B-Biological Sci.*, 370.

Brockington, D. & Wilkie, D. (2015b). Protected areas and poverty. *Phil. Trans. R. Soc. B*, 370, 20140271.

Brun, C., Cook, A.R., Lee, J.S.H., Wich, S.A., Koh, L.P. & Carrasco, L.R. (2015). Analysis of deforestation and protected area effectiveness in Indonesia: A comparison of Bayesian spatial models. *Glob. Environ. Chang.*, 31, 285–295.

Bruner, A.G., Gullison, R.E., Rice, R.E. & da Fonseca, G.A.B. (2001). Effectiveness of parks in protecting tropical biodiversity. *Science (80-.).*, 291, 125–128.

Busch, J., Engelmann, J., Cook-Patton, S.C., Griscom, B.W., Kroeger, T., Possingham, H. & Shyamsundar, P. (2019). Potential for low-cost carbon dioxide removal through tropical reforestation. *Nat. Clim. Chang.*

Büscher, B., Fletcher, R., Brockington, D., Sandbrook, C., Adams, W.M., Campbell, L., Corson, C., Dressler, W., Duffy, R., Gray, N., Holmes, G., Kelly, A., Lunstrum, E., Ramutsindela, M. & Shanker, K. (2017). Half-Earth or Whole Earth? Radical ideas for conservation, and their implications. *ORYX*, 51, 407–410.

Canavire-Bacarreza, G. & Hanauer, M.M. (2013). Estimating the Impacts of Bolivia's Protected Areas on Poverty. *World Dev.*, 41, 265–285.

Cardinale, B.J., Duffy, J.E., Gonzalez, A., Hooper, D.U., Perrings, C., Venail, P., Narwani, A., Mace, G.M., Tilman, D., Wardle, D.A., Kinzig, A.P., Daily, G.C., Loreau, M., Grace, J.B., Larigauderie, A., Srivastava, D.S. & Naeem, S. (2012). Biodiversity loss and its impact on humanity. *Nature*, 486, 59–67.

Carranza, T., Balmford, A., Kapos, V. & Manica, A. (2014). Protected area effectiveness in reducing conversion in a rapidly vanishing ecosystem: The Brazilian Cerrado. *Conserv. Lett.*, 7, 216–223.

Cernea, M.M. & Schmidt-Soltau, K. (2006). Poverty risks and national parks: Policy issues in conservation and resettlement. *World Dev.*, 34, 1808–1830.

Chen, S. & Ravallion, M. (2008). The developing world is poorer than we thought, but no less successful in the fight against poverty. *World Bank Policy Res. Work. Pap. Ser. Vol*.

Chen, X. & Nordhaus, W.D. (2011). Using luminosity data as a proxy for economic statistics. *Proc. Natl. Acad. Sci. U. S. A.*, 108, 8589–8594.

Christie, A.P., Amano, T., Martin, P.A., Shackelford, G.E., Simmons, B.I. & Sutherland, W.J. (2019). Simple study designs in ecology produce inaccurate estimates of biodiversity responses. *J. Appl. Ecol.*, 1365–2664.13499.

Coad, L., Watson, J.E.M., Geldmann, J., Burgess, N.D., Leverington, F., Hockings, M., Knights, K. & Di Marco, M. (2019). Widespread shortfalls in protected area resourcing undermine efforts to conserve biodiversity. *Front. Ecol. Environ.*, 17, 259–264.

Coetzee, B.W.T., Gaston, K.J. & Chown, S.L. (2014). Local scale comparisons of biodiversity as a test for global protected area ecological performance: A meta-analysis. *PLoS One*, 9.

Coria, J. & Calfucura, E. (2012). Ecotourism and the development of indigenous communities: The good, the bad, and the ugly. *Ecol. Econ.*, 73, 47–55.

Costanza, R., D'Arge, R., De Groot, R., Farber, S., Grasso, M., Hannon, B., Limburg, K., Naeem, S., O'Neill, R. V., Paruelo, J., Raskin, R.G., Sutton, P. & Van Den Belt, M. (1997). The value of the world's ecosystem services and natural capital. *Nature*, 387, 253–260.

Craigie, I.D., Baillie, J.E.M., Balmford, A., Carbone, C., Collen, B., Green, R.E. & Hutton, J.M. (2010). Large mammal population declines in Africa's protected areas. *Biol. Conserv.*, 143, 2221–2228.

Curran, L.M., Trigg, S.N., McDonald, A.K., Astiani, D., Hardiono, Y.M., Siregar, P., Caniago, I. & Kasischke, E. (2004). Lowland forest loss in protected areas of Indonesian Borneo. *Science (80-.).*, 303, 1000–1003.

Deaton, A. (2005). Measuring Poverty in a Growing World (or Measuring Growth in a Poor World). *Rev. Econ. Stat.*, 87, 1–19.

Deaton, A.S. (2009). *Instruments of development: Randomization in the tropics, and the search for the elusive keys to economic development*.

Doherty, T.S., Glen, A.S., Nimmo, D.G., Ritchie, E.G. & Dickman, C.R. (2016). Invasive predators and global biodiversity loss. Proc. Natl. Acad. Sci. U. S. A., 113, 11261–11265.

Dudley, N. (2008). *Guidelines for applying protected area management categories*. IUCN.

Edwards, D.P., Socolar, J.B., Mills, S.C., Burivalova, Z., Koh, L.P. & Wilcove, D.S. (2019). Conservation of Tropical Forests in the Anthropocene. Curr. Biol.

Edwards, D.P., Tobias, J.A., Sheil, D., Meijaard, E. & Laurance, W.F. (2014). Maintaining ecosystem function and services in logged tropical forests. Trends Ecol. Evol.

Engert, M., Warren, T. & Baker, K.J.M. (2019). Germany Has Frozen Funding For Wildlife Charity WWF Amid Ongoing Human Rights Investigations. *BuzzFeed News*.

Ferraro, P.J. & Hanauer, M.M. (2014). Quantifying causal mechanisms to determine how protected areas affect poverty through changes in ecosystem services and infrastructure. *Proc. Natl. Acad. Sci.*, 111.

Ferraro, P.J. & Hanauer, M.M. (2015). Through what mechanisms do protected areas affect environmental and social outcomes? *Philos. Trans. R. Soc. B-Biological Sci.*, 370.

Ferraro, P.J., Hanauer, M.M., Miteva, D.A., Javier Canavire-Bacarreza, G., Pattanayak, S.K. & Sims, K.R.E. (2013). More strictly protected areas are not necessarily more protective: evidence from Bolivia, Costa Rica, Indonesia, and Thailand. *Environ. Res. Lett.*, 8.

Ferraro, P.J., Sanchirico, J.N. & Smith, M.D. (2018). Causal inference in coupled human and natural systems. *Proc. Natl. Acad. Sci. U. S. A.*, 201805563.

Ferreira, F.H.G., Leite, P.G. & Ravallion, M. (2010). Poverty reduction without economic growth?: Explaining Brazil's poverty dynamics, 1985–2004. *J. Dev. Econ.*, 93, 20–36.

Fisher, B., Balmford, A., Ferraro, P.J., Glew, L., Mascia, M., Naidoo, R. & Ricketts, T.H. (2014). Moving Rio forward and avoiding 10 more years with little evidence for effective conservation policy. *Conserv. Biol.*, 28, 880–882.

Fisher, B. & Christopher, T. (2007). Poverty and biodiversity: measuring the overlap of human poverty and the biodiversity hotspots. *Ecol. Econ.*, 62, 93–101.

Foley, J.A., Defries, R., Asner, G.P., Barford, C., Bonan, G., Carpenter, S.R., Chapin, F.S., Coe, M.T., Daily, G.C., Gibbs, H.K., Helkowski, J.H., Holloway, T., Howard, E.A., Kucharik, C.J., Monfreda, C., Patz, J.A., Prentice, I.C., Ramankutty, N. & Snyder, P.K. (2005). Global consequences of land use. *Science*, 309, 570–4.

Geldmann, J., Barnes, M., Coad, L., Craigie, I.D., Hockings, M. & Burgess, N.D. (2013). Effectiveness of terrestrial protected areas in reducing habitat loss and population declines. *Biol. Conserv.*, 161, 230–238.

Golden, C.D., Fernald, L.C.H., Brashares, J.S., Rasolofoniaina, B.J.R. & Kremen, C. (2011). Benefits of wildlife consumption to child nutrition in a biodiversity hotspot. Proc. Natl. Acad. Sci. U. S. A., 108, 19653–19656.

Golden Kroner, R.E., Qin, S., Cook, C.N., Krithivasan, R., Pack, S.M., Bonilla, O.D., Cort-Kansinally, K.A., Coutinho, B., Feng, M., Martínez Garcia, M.I., He, Y., Kennedy, C.J., Lebreton, C., Ledezma, J.C., Lovejoy, T.E., Luther, D.A., Parmanand, Y., Ruíz-Agudelo, C.A., Yerena, E., Morón Zambrano, V. & Mascia, M.B. (2019). The uncertain future of protected lands and waters. *Science*, 364, 881–886.

Gonner, C., Center for International Forestry Research, J. (Indonesia) eng, Haug, M., Cahyat, A., Wollenberg, E., DeJong, W., Limberg, G., Cronkleton, P., Moeliono, M. & Becker, M. (2007). Capturing nested spheres of poverty a model for multidimensional povery analysis and monitoring.

Gray, C.L., Hill, S.L.L., Newbold, T., Hudson, L.N., Börger, L., Contu, S., Hoskins, A.J., Ferrier, S., Purvis, A. & Scharlemann, J.P.W. (2016). Local biodiversity is higher inside than outside terrestrial protected areas worldwide. *Nat. Commun.*, 7, 12306.

Griggs, D., Stafford-Smith, M., Gaffney, O., Rockström, J., Öhman, M.C., Shyamsundar, P., Steffen, W., Glaser, G., Kanie, N. & Noble, I. (2013). Sustainable development goals for people and planet. *Nature*, 495, 305–307.

Grosh, M.E. & Glewwe, P. (1998). Data Watch: The World Bank's Living Standards Measurement Study Household Surveys Margaret E. Grosh and Paul Glewwe. *J. Econ. Perspect.*, 12, 187–196.

Hanauer, M.M. & Canavire-Bacarreza, G. (2015). Implications of heterogeneous impacts of protected areas on deforestation and poverty. *Philos. Trans. R. Soc. B-Biological Sci.*, 370.

Hansen, M.C., Potapov, P. V, Moore, R., Hancher, M., Turubanova, S.A., Tyukavina, A., Thau, D., Stehman, S. V, Goetz, S.J., Loveland, T.R., Kommareddy, A., Egorov, A., Chini, L., Justice, C.O. & Townshend, J.R.G. (2013). High-Resolution Global Maps of 21st-Century Forest Cover Change. *Science (80-.).*, 342, 850–853.

Hauenstein, S., Kshatriya, M., Blanc, J., Dormann, C.F. & Beale, C.M. (2019). African elephant poaching rates correlate with local poverty, national corruption and global ivory price. *Nat. Commun.*, 10, 2242.

Heinen, J.T. & Shrestha, S.K. (2006). Evolving policies for conservation: an historical profile of the protected area system of Nepal. *J. Environ. Plan. Manag.*, 49, 41–58.

Heino, M., Kummu, M., Makkonen, M., Mulligan, M., Verburg, P.H., Jalava, M. & Räsänen, T.A. (2015). Forest loss in protected areas and intact forest landscapes: A global analysis. *PLoS One*, 10.

Herrera, D., Pfaff, A. & Robalino, J. (2019). Impacts of protected areas vary with the level of government: Comparing avoided deforestation across agencies in the Brazilian Amazon. *Proc. Natl. Acad. Sci. U. S. A.*, 116, 14916–14925.

Holmes, G. (2007). Protection, politics and protest: understanding resistance to conservation. *Conserv. Soc.*, 5, 184.

Isbell, F., Craven, D., Connolly, J., Loreau, M., Schmid, B., Beierkuhnlein, C., Bezemer, T.M., Bonin, C., Bruelheide, H., De Luca, E., Ebeling, A., Griffin, J.N., Guo, Q., Hautier, Y., Hector, A., Jentsch, A., Kreyling, J., Lanta, V., Manning, P., Meyer, S.T., Mori, A.S., Naeem, S., Niklaus, P.A., Polley, H.W., Reich, P.B., Roscher, C., Seabloom, E.W., Smith, M.D., Thakur, M.P., Tilman, D., Tracy, B.F., Van Der Putten, W.H., Van Ruijven, J., Weigelt, A., Weisser, W.W., Wilsey, B. & Eisenhauer, N. (2015). Biodiversity increases the resistance of ecosystem productivity to climate extremes. *Nature*, 526, 574–577.

Jagger, P., Babigumira, R., Hogarth, N.J., Bauch, S., Börner, J. & Smith-Hall, C. (2014). Environmental Income and Rural Livelihoods: A Global-Comparative Analysis. *World Dev.*, 64, S12–S28.

Jean, N., Burke, M., Xie, M., Davis, W.M., Lobell, D.B. & Ermon, S. (2016). Combining satellite imagery and machine learning to predict poverty. *Science (80-.).*, 353, 790–794.

Joppa, L.N. & Pfaff, A. (2009). High and Far: Biases in the Location of Protected Areas. *PLoS One*, 4.

Karanth, K.K. (2007). Making resettlement work: The case of India's Bhadra wildlife sanctuary. *Biol. Conserv.*, 139, 315–324.

Kuempel, C.D., Adams, V.M., Possingham, H.P. & Bode, M. (2018). Bigger or better: The relative benefits of protected area network expansion and enforcement for the conservation of an exploited species. *Conserv. Lett.*, 11, e12433.

Lambin, E.F. & Meyfroidt, P. (2011). Global land use change, economic globalization, and the looming land scarcity. *Proc. Natl. Acad. Sci. U. S. A.*, 108, 3465–72.

Lambin, E.F., Meyfroidt, P., Rueda, X., Blackman, A., Börner, J., Cerutti, P.O., Dietsch, T., Jungmann, L., Lamarque, P., Lister, J., Walker, N.F. & Wunder, S. (2014). Effectiveness and synergies of policy instruments for land use governance in tropical regions. *Glob. Environ. Chang.*, 28, 129–140.

Laurance, W.F., Useche, D.C., Rendeiro, J., Kalka, M., Bradshaw, C.J.A., Sloan, S.P., Laurance, S.G., Campbell, M., Abernethy, K., Alvarez, P., Arroyo-Rodriguez, V., Ashton, P., Benitez-Malvido, J., Blom, A., Bobo, K.S., Cannon, C.H., Cao, M., Carroll, R., Chapman, C., Coates, R., Cords, M., Danielsen, F., De Dijn, B., Dinerstein, E., Donnelly, M.A., Edwards, D., Edwards, F., Farwig, N., Fashing, P., Forget, P.-M., Foster, M., Gale, G., Harris, D., Harrison, R., Hart, J., Karpanty, S., Kress, W.J., Krishnaswamy, J., Logsdon, W., Lovett, J., Magnusson, W., Maisels, F., Marshall, A.R., McClearn, D., Mudappa, D., Nielsen, M.R., Pearson, R., Pitman, N., van der Ploeg, J., Plumptre, A., Poulsen, J., Quesada, M., Rainey, H., Robinson, D., Roetgers, C., Rovero, F., Scatena, F., Schulze, C., Sheil, D., Struhsaker, T., Terborgh, J., Thomas, D., Timm, R., Urbina-Cardona, J.N., Vasudevan, K., Wright, S.J., Arias-G, J.C., Arroyo, L., Ashton, M., Auzel, P., Babaasa, D., Babweteera, F., Baker, P., Banki, O., Bass, M., Bila-Isia, I., Blake, S., Brockelman, W., Brokaw, N., Bruehl, C.A., Bunyavejchewin, S., Chao, J.-T., Chave, J., Chellam, R., Clark, C.J., Clavijo, J., Congdon, R., Corlett, R., Dattaraja, H.S., Dave, C., Davies, G., Beisiegel, B. de M., da Silva, R. de N., Di Fiore, A., Diesmos, A., Dirzo, R., Doran-Sheehy, D., Eaton, M., Emmons, L., Estrada, A. & others. (2012). Averting biodiversity collapse in tropical forest protected areas. *Nature*, 489, 290–+.

Lecun, Y., Bottou, L., Bengio, Y. & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proc. IEEE*, 86, 2278–2324.

Lipton, M. & Ravallion, M. (1995). Poverty and policy. *Handb. Dev. Econ.*, 3, 2551–2657.

Liu, J., Dietz, T., Carpenter, S.R., Folke, C., Alberti, M., Redman, C.L., Schneider, S.H., Ostrom, E., Pell, A.N., Lubchenco, J., Taylor, W.W., Ouyang, Z., Deadman, P., Kratz, T. & Provencher, W. (2007). Coupled human and natural systems. *Ambio*.

Lu, Y., Nakicenovic, N., Visbeck, M. & Stevance, A.S. (2015). Policy: Five priorities for the un Sustainable Development Goals. *Nature*.

Ma, B., Cai, Z., Zheng, J. & Wen, Y. (2019). Conservation, ecotourism, poverty, and income inequality – A case study of nature reserves in Qinling, China. World Dev., 115, 236–244.Marvier, M. (2014). New Conservation Is True Conservation. *Conserv. Biol.*, 28, 1–3.

Maldonado, J.H., Moreno-Sánchez, R. del P., Espinoza, S., Bruner, A., Garzón, N. & Myers, J. (2018). Peace is much more than doves: The economic benefits of bird-based tourism as a result of the peace treaty in Colombia. World Dev., 106, 78–86.

Mather, A.S. (1992). The forest transition. *Area*, 24, 367–379.

De Merode, E., Homewood, K. & Cowlishaw, G. (2004). The value of bushmeat and other wild foods to rural households living in extreme poverty in Democratic Republic of Congo. Biol. Conserv., 118, 573–581.

Meyfroidt, P. & Lambin, E.F. (2011). Global Forest Transition: Prospects for an End to Deforestation. *Annu. Rev. Environ. Resour.*, 36, 343–371.

Meyfroidt, P., Roy Chowdhury, R., de Bremond, A., Ellis, E.C., Erb, K.-H., Filatova, T., Garrett, R.D., Grove, J.M., Heinimann, A., Kuemmerle, T., Kull, C.A., Lambin, E.F., Landon, Y., le Polain de Waroux, Y., Messerli, P., Müller, D., Nielsen, J.Ø., Peterson, G.D., Rodriguez García, V., Schlüter, M., Turner, B.L. & Verburg, P.H. (2018). Middle-range theories of land system change. *Glob. Environ. Chang.*, 53, 52–67.

Miranda, J.J., Corral, L., Blackman, A., Asner, G. & Lima, E. (2014). *Effects of Protected Areas on Forest Cover Change and Local Communities: Evidence from the Peruvian Amazon*.

Miteva, D.A., Pattanayak, S.K. & Ferraro, P.J. (2012). Evaluation of biodiversity policy instruments: what works and what doesn't? *Oxford Rev. Econ. Policy*, 28, 69–92.

Naidoo, R., Gerkey, D., Hole, D., Pfaff, A., Ellis, A.M., Golden, C.D., Herrera, D., Johnson, K., Mulligan, M., Ricketts, T.H. & Fisher, B. (2019). Evaluating the impacts of protected areas on human well-being across the developing world. *Sci. Adv.*, 5, eaav3006.

Narayan, D. (2000). Voices of the poor: can anyone hear us? Voices poor can anyone Hear us?

Naughton-Treves, L., Holland, M.B. & Brandon, K. (2005). The role of protected areas in conserving biodiversity and sustaining local livelihoods. *Annu. Rev. Environ. Resour.*, 30, 219–252.

Nelson, A. & Chomitz, K.M. (2011). Effectiveness of Strict vs. Multiple Use Protected Areas in Reducing Tropical Forest Fires: A Global Analysis Using Matching Methods. *PLoS One*, 6.

Newbold, T. (2018). Future effects of climate and land-use change on terrestrial vertebrate community diversity under different scenarios. In: Proc. R. Soc. B Biol. Sci. Royal Society Publishing.

Newbold, T., Hudson, L.N., Hill, S.L.L., Contu, S., Lysenko, I., Senior, R.A., Börger, L., Bennett, D.J., Choimes, A., Collen, B., Day, J., De Palma, A., Díaz, S., Echeverria-Londoño, S., Edgar, M.J., Feldman, A., Garon, M., Harrison, M.L.K., Alhusseini, T., Ingram, D.J., Itescu, Y., Kattge, J., Kemp, V., Kirkpatrick, L., Kleyer, M., Correia, D.L.P., Martin, C.D., Meiri, S., Novosolov, M., Pan, Y., Phillips, H.R.P., Purves, D.W., Robinson, A., Simpson, J., Tuck, S.L., Weiher, E., White, H.J., Ewers, R.M., MacE, G.M., Scharlemann, J.P.W. & Purvis, A. (2015). Global effects of land use on local terrestrial biodiversity. *Nature*, 520, 45–50.

Noor, A.M., Alegana, V.A., Gething, P.W., Tatem, A.J. & Snow, R.W. (2008). Using remotely sensed night-time light as a proxy for poverty in Africa. *Popul. Health Metr.*, 6, 5.

Oates, J.F. (1999). *Myth and reality in the rain forest: How conservation strategies are failing in West Africa*. Univ of California Press.

Ostrom, E. (2009). A general framework for analyzing sustainability of social-ecological systems. *Science (80-.).*

Ostrom, E. & Nagendra, H. (2006). Insights on linking forests, trees, and people from the air, on the ground, and in the laboratory. *Proc. Natl. Acad. Sci. U. S. A.*

Pack, S.M., Ferreira, M.N., Krithivasan, R., Murrow, J., Bernard, E. & Mascia, M.B. (2016). Protected area downgrading, downsizing, and degazettement (PADDD) in the Amazon. *Biol. Conserv.*, 197, 32–39.

Pfeifer, M., Lefebvre, V., Peres, C.A., Banks-Leite, C., Wearn, O.R., Marsh, C.J., Butchart, S.H.M., Arroyo-Rodríguez, V., Barlow, J., Cerezo, A., Cisneros, L., D'Cruze, N., Faria, D., Hadley, A., Harris, S.M., Klingbeil, B.T., Kormann, U., Lens, L., Medina-Rangel, G.F., Morante-Filho, J.C., Olivier, P., Peters, S.L., Pidgeon, A., Ribeiro, D.B., Scherber, C., Schneider-Maunoury, L., Struebig, M., Urbina-Cardona, N., Watling, J.I., Willig, M.R., Wood, E.M. & Ewers, R.M. (2017). Creation of forest edges has a global impact on forest vertebrates. Nature, 551, 187–191.

Pimm, S.L., Jenkins, C.N., Abell, R., Brooks, T.M., Gittleman, J.L., Joppa, L.N., Raven, P.H., Roberts, C.M. & Sexton, J.O. (2014). The biodiversity of species and their rates of extinction, distribution, and protection. *Science (80-.).*

Powers, R.P. & Jetz, W. (2019). Global habitat loss and extinction risk of terrestrial vertebrates under future land-use-change scenarios. Nat. Clim. Chang.

Pringle, R.M. (2017). Upgrading protected areas to conserve wild biodiversity. *Nature*, 546, 91–99.

Pullin, A.S., Bangpan, M., Dalrymple, S., Dickson, K., Haddaway, N.R., Healey, J.R., Hauari, H., Hockley, N., Jones, J.P.G. & Knight, T. (2013). Human well-being impacts of terrestrial protected areas. *Environ. Evid.*, 2, 19.

Pyhälä, A.A., Oroza, A.O. & Counsell, S. (2016). Protected Areas in the Congo Basin: Failing both People and Biodiversity?

Rasmussen, L.V., Coolsaet, B., Martin, A., Mertz, O., Pascual, U., Corbera, E., Dawson, N., Fisher, J.A., Franks, P. & Ryan, C.M. (2018). Social-ecological outcomes of agricultural intensification. *Nat. Sustain.*, 1, 275–282.

Ravallion, M. (1998). *Poverty lines in theory and practice*. World Bank Publications.

Ravallion, M. (2001). Growth, inequality and poverty: Looking beyond averages. *World Dev.*, 29, 1803–1815.

Ricardo, D. (1891). *Principles of political economy and taxation*. G. Bell.

Robinson, W.C. (2002). Population policy in early Victorian England. *Eur. J. Popul.*, 18, 153–173.

Rockström, J., Steffen, W., Noone, K., Persson, Å., Chapin, F.S., Lambin, E.F., Lenton, T.M., Scheffer, M., Folke, C., Schellnhuber, H.J., Nykvist, B., de Wit, C.A., Hughes, T., van der Leeuw, S., Rodhe, H., Sörlin, S., Snyder, P.K., Costanza, R., Svedin, U., Falkenmark, M., Karlberg, L., Corell, R.W., Fabry, V.J., Hansen, J., Walker, B., Liverman, D., Richardson, K., Crutzen, P. & Foley, J.A. (2009). A safe operating space for humanity. *Nature*, 461, 472–475.

Rodrigues, A.S.L., Andelman, S.J., Bakarr, M.I., Boitani, L., Brooks, T.M., Cowling, R.M., Fishpool, L.D.C., da Fonseca, G.A.B., Gaston, K.J., Hoffmann, M., Long, J.S., Marquet, P.A., Pilgrim, J.D., Pressey, R.L., Schipper, J., Sechrest, W., Stuart, S.N., Underhill, L.G., Waller, R.W., Watts, M.E.J. & Yan, X. (2004). Effectiveness of the global protected area network in representing species diversity. *Nature*, 428, 640–643.

Roe, D. (2019). Biodiversity loss—more than an environmental emergency. *Lancet Planet. Heal.*

Scharlemann, J.P.W., Kapos, V., Campbell, A., Lysenko, I., Burgess, N.D., Hansen, M.C., Gibbs, H.K., Dickson, B. & Miles, L. (2010). Securing tropical forest carbon: The contribution of protected areas to REDD. *ORYX*, 44, 352–357.

Scheffers, B.R., Oliveira, B.F., Lamb, I. & Edwards, D.P. (2019). Global wildlife trade across the tree of life. Science (80-.)., 366, 71–76.

Scoones, I. (1998). Sustainable Rural Livelihoods: A Framework for Analysis. IDS.

Senior, R.A., Hill, J.K. & Edwards, D.P. (2019). Global loss of climate connectivity in tropical forests. Nat. Clim. Chang.

Shah, P. & Baylis, K. (2015). Evaluating heterogeneous conservation effects of forest protection in Indonesia. PLoS One, 10.

Sims, K.R.E. & Alix-Garcia, J.M. (2017). Parks versus PES: Evaluating direct and incentive-based land conservation in Mexico. *J. Environ. Econ. Manage.*, 86, 8–28.

Song, X.P., Hansen, M.C., Stehman, S. V., Potapov, P. V., Tyukavina, A., Vermote, E.F. & Townshend, J.R. (2018). Global land change from 1982 to 2016. *Nature*.

Soule, M. (2013). The "New Conservation." *Conserv. Biol.*, 27, 895–897.

Spatz, D.R., Zilliacus, K.M., Holmes, N.D., Butchart, S.H.M., Genovesi, P., Ceballos, G., Tershy, B.R. & Croll, D.A. (2017). Globally threatened vertebrates on islands with invasive species. Sci. Adv., 3, e1603080.

Stiglitz, J.E., Sen, A. & Fitoussi, J.-P. (2017). Report by the commission on the measurement of economic performance and social progress.

Stuart, E.A. (2010). Matching Methods for Causal Inference: A Review and a Look Forward. *Stat. Sci.*, 25, 1–21.

Symes, W.S., Edwards, D.P., Miettinen, J., Rheindt, F.E. & Carrasco, L.R. (2018). Combined impacts of deforestation and wildlife trade on tropical biodiversity are severely underestimated. *Nat. Commun.*, 9.

Von Thünen, J.H. (1966). *Isolated state: an English edition of Der isolierte Staat*. Pergamon Press.

Tittensor, D.P., Walpole, M., Hill, S.L.L., Boyce, D.G., Britten, G.L., Burgess, N.D., Butchart, S.H.M., Leadley, P.W., Regan, E.C., Alkemade, R., Baumung, R., Bellard, C., Bouwman, L., Bowles-Newark, N.J., Chenery, A.M., Cheung, W.W.L., Christensen, V., Cooper, H.D., Crowther, A.R., Dixon, M.J.R., Galli, A., Gaveau, V., Gregory, R.D., Gutierrez, N.L., Hirsch, T.L., Hoeft, R., Januchowski-Hartley, S.R., Karmann, M., Krug, C.B., Leverington, F.J., Loh, J., Lojenga, R.K., Malsch, K., Marques, A., Morgan, D.H.W., Mumby, P.J., Newbold, T., Noonan-Mooney, K., Pagad, S.N., Parks, B.C., Pereira, H.M., Robertson, T., Rondinini, C., Santini, L., Scharlemann, J.P.W., Schindler, S., Sumaila, U.R., Teh, L.S.L., van Kolck, J., Visconti, P. & Ye, Y. (2014). A mid-term analysis of progress toward international biodiversity targets. *Science (80-.).*, 346, 241–244.

Venter, O., Fuller, R.A., Segan, D.B., Carwardine, J., Brooks, T., Butchart, S.H.M., Di Marco, M., Iwamura, T., Joseph, L., O'Grady, D., Possingham, H.P., Rondinini, C., Smith, R.J., Venter, M. & Watson, J.E.M. (2014). Targeting Global Protected Area

Expansion for Imperiled Biodiversity. *Plos Biol.*, 12.

Waldron, A., Miller, D.C., Redding, D., Mooers, A., Kuhn, T.S., Nibbelink, N., Roberts, J.T., Tobias, J.A. & Gittleman, J.L. (2017). Reductions in global biodiversity loss predicted from conservation spending. *Nature*, 551, 364–367.

Walpole, M.J. & Leader-Williams, N. (2001). Masai Mara tourism reveals partnership benefits. *Nature*, 413, 771.

Warren, T. & Baker, K.J.M. (2019). WWF Funds Guards Who Have Tortured And Killed People. *BuzzFeed News*.

Watmough, G.R., Atkinson, P.M., Saikia, A. & Hutton, C.W. (2016). Understanding the Evidence Base for Poverty–Environment Relationships using Remotely Sensed Satellite Data: An Example from Assam, India. *World Dev.*, 78, 188–203.

Watmough, G.R., Marcinko, C.L.J., Sullivan, C., Tschirhart, K., Mutuo, P.K., Palm, C.A. & Svenning, J.-C. (2019). Socioecologically informed use of remote sensing data to predict rural household poverty. *Proc. Natl. Acad. Sci.*, 116, 1213–1218.

Watson, J.E.M., Dudley, N., Segan, D.B. & Hockings, M. (2014). The performance and potential of protected areas. *Nature*, 515, 67–73.

Weber, J.G., Sills, E.O., Bauch, S. & Pattanayak, S.K. (2011). Do ICDPs Work? An Empirical Evaluation of Forest-Based Microenterprises in the Brazilian Amazon. *Land Econ.*, 87, 661–681.

West, P., Igoe, J. & Brockington, D. (2006). Parks and peoples: The social impact of protected areas. *Annu. Rev. Anthropol.*, 35, 251–277.

Wunder, S. (2000). Ecotourism and economic incentives - an empirical approach. Ecol. Econ., 32, 465–479.

Chapter 2:

Impact of Protected Areas on Poverty, Extreme Poverty and Inequality in Nepal

This chapter has been published as:

Den Braber, B., Evans, K.L., and Oldekop, J.A. (2018) Impacts of protected areas on poverty, extreme poverty and inequality in Nepal. *Conservation Letters*, **11**, e12576.

K.L.E. and J.A.O. are joint last authors.

2.1 Abstract

Protected areas (PAs) are key for biodiversity conservation, but there are concerns that they can exacerbate poverty or unequal access to potential benefits, such as those arising from tourism. We assess how Nepalese PAs influence poverty, extreme poverty and inequality using a multidimensional poverty index, and a quasi-experimental design that controls for potential confounding factors in non-random treatment allocation. We specifically investigate the role of tourism in contributing to PA impacts. Nepali PAs reduced overall poverty and extreme poverty, and crucially, did not exacerbate inequality. Benefits occurred in lowland and highland regions, and were often greater when a larger proportion of the area was protected. Spread of benefits to nearby areas outside PAs was negligible. Furthermore, older PAs performed better than more recently established ones, suggesting the existence of time-lags. Although tourism was a key driver of poverty alleviation, PAs also reduced extreme poverty in areas with fewer tourists.

2.2 Introduction

Protected areas (PAs) are key conservation strategies but also have socioeconomic impacts on people living in and around them (Brockington and Wilkie 2015). PAs limiting anthropogenic activities can harm local economic development (Brockington and Wilkie 2015), but can also safeguard ecosystem services that local communities depend on, and generate additional sources of income, for example through tourism (Ferraro and Hanauer 2015). Some studies find that PAs are linked to high poverty levels (de Sherbinin 2008; Fisher and Christopher 2007), but such associations can be confounded because PAs are often located in areas with limited development potential (Joppa and Pfaff 2009). There are, therefore, growing efforts to assess PA outcomes using techniques that control for this non-random allocation of PAs. Such studies provide increasing evidence that PAs can reduce poverty, albeit with much heterogeneity in effect sizes (Andam *et al.* 2010; Hanauer and Canavire-Bacarreza 2015; Miranda *et al.* 2014; Sims and Alix-Garcia 2016; Yergeau *et al.* 2017).

Despite this progress, several topics remain understudied. First, PA assessments have focused primarily on mean poverty outcomes across entire communities (but see Sims 2010). However, PA's financial benefits may suffer from elite capture (Agrawal and Gupta 2005), leading to greater inequalities. Assessing the mechanisms through which PAs influence poverty is essential. PAs may increase tourism opportunities leading to improved local income and employment (Walpole and Leader-Williams 2001). Assessing tourism impacts across large spatial extents is often limited by data availability, and assessments have predominantly used binary proxies (presence or absence of tourism infrastructure, Ferraro *et al.* 2011). Yet to gain a better understanding of how tourism contributes to local poverty alleviation, it is important to move beyond binary assessments of tourism and consider variation in the intensity of tourism in PAs (Robalino and Villalobos 2015). Finally, there is substantial spatial variation in the proportion of land surrounding a community that is protected, the duration that it has been protected for, and livelihood opportunities that are constrained by a series of factors, such as slope and elevation that influence agricultural suitability (Gentle and Maraseni 2012). These factors can influence the magnitude, and possibly even the direction of PA effects on poverty.

Here, we assess how PAs in Nepal influence multiple measures of poverty. We combine national census derived poverty estimates for 2001 and 2011, and use statistical matching to construct a counterfactual group. We build upon previous research by i) quantifying how PA status influences measures of extreme poverty and inequality, in addition to overall measures of poverty, ii) using tourism indicators to assess if tourism is an important mechanism through which PAs influence poverty, and iii) testing whether effects of PAs on poverty are moderated by variations in the amount of protected land, time since establishment, and elevation.

Nepal provides a good case study to assess the effects of protected areas on multiple poverty outcomes. It is one of the poorest Asian countries (Alkire and Santos 2011) and has an extensive PA network, covering 20% of the country’s land surface. Nepalese PA policies were first characterized by a strict “fences and fines” approach (Heinen and Shrestha 2006) that denied local people’s user rights. However, during the nineties several important pieces of legislation were passed to promote social welfare including redistribution initiatives to minimize inequality by spending 30-50% of PA revenues on community development (Spiteri and Nepal 2008).

2.3 Methods

2.3.1 Data

We compiled a high spatial-resolution, national-level dataset using 3,845 of Nepal’s 3,973 Village Development Committees (VDCs), the sub-district level administrative unit, as our unit of analysis.

i) Poverty metrics

We use household health, education and living standards data from the Nepali national censuses of 2001 and 2011 to develop three multi-dimensional poverty (MDP) measures based on the multi-dimensional poverty index developed by Alkire and Santos (2011): poverty (MDP>0.33 - following Alkire and Santos' (2011) cut-off for measuring poverty); extreme poverty (MDP>0.66 - this doubles the standard poverty threshold, following other studies (e.g. Lokshin and Ravallion, 2000) and indicates that at a minimum a household is completely deprived in one of the three poverty dimensions and partially deprived across the remaining two dimensions); and inequality - measured as the standard deviation of the incidence of household poverty (S2.1; Figure 2.1A). Using alternative thresholds for defining extreme poverty either generates too few VDCs that contain extreme poverty (70% threshold – 314 VDCs using 2001 baseline data compared to 1,153 with a 66% threshold) or generates qualitatively identical results and conclusions (60% threshold, Figure S2.2).

ii) Defining Protected Area treatments

We define protected treatments as VDCs that overlap Nepal's 32 PAs (IUCN categories II - VI, Nepal lacks category I PAs) using the World Database on Protected Areas (WDPA; IUCN & UNEP-WCMC, 2016; Figure 2.1B). The vast majority of these are multiple-use PAs. We conduct two separate analyses: one focusing on PAs established before 2001 (the baseline year of our poverty data), and one focusing on PAs established between 2001 and 2011. We conduct this second analysis as a robustness check because PAs established prior to 2001 could affect our baseline measures, although baseline poverty metrics were similar in VDCs that were protected before and after 2001 (see Fig. 2.2). We also defined protected VDCs using two separate definitions: those with i) at least 10% of their area overlapping with a PA (e.g. Andam *et al.* 2010; Hanauer and Canavire-Bacarreza, 2015) and ii) at least 70% of the VDC being protected (which is close to the mean percentage overlap for overlapping VDCs - PAs established before 2001 = 65.2%; PAs established after 2011 = 71.4%). VDCs with <1% of their area protected were defined as non-protected to ensure a clear distinction in the magnitude of protection between control and treatment VDCs.

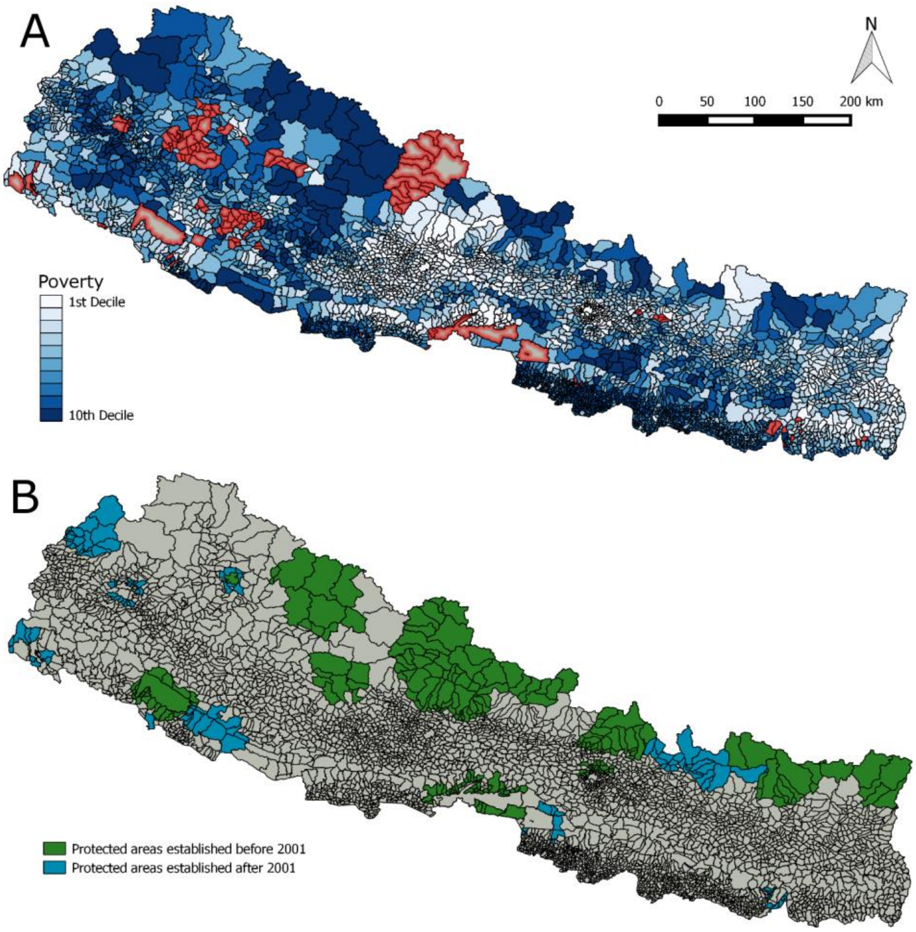

Figure 2.1. Poverty and protected areas. (A) Multidimensional Poverty in 2011. Each polygon represents a Village Development Committee (VDC). Data are presented as deciles. Grey areas with red contours represent excluded VDCs (reasons for exclusion include missing data due to armed conflict and instances of inconsistent data from the Nepali Department of Forests). (B) Schematic map of protected areas in Nepal. Data from the World Database of Protected Areas. In our analysis we included 192 VDCs that were protected before 2001 (of which 110 were protected using the 70% threshold definition), and 106 VDCs that were protected between 2001 and 2011 (of which 67 were protected using the 70% threshold).

iii) Tourism metrics

We assessed how PAs with different tourism intensities impacted our outcome variables, using data on official tourism numbers for each PA in 2011 (low < 10,000 visitors, intermediate 10,000-100,000, high > 100,000; Nepal Tourism Statistics, 2013). We also assessed how proximity to a PA entrance and trekking routes (categorised as major or minor; Table S2.9) contributed to heterogeneity in PA impacts using a mean travel time estimate (weighted by population density) from each VDC to the nearest PA entrance, and major and minor trekking routes (S2.3).

iv) Confounding factors

We selected a suite of biophysical and socioeconomic covariates based on their potential to influence the outcome or the relationship between treatment and outcomes. These covariates were baseline levels of our poverty measures, slope, elevation, precipitation, VDC area, forest cover, travel time from the VDC to population centres and district headquarters, proportion of the VDC under community forest management and the age of community forestry arrangement, population density, agricultural effort, international migration and district (Table S2.3).

2.3.2 Matching and post-matching analyses

We used a combined matching and regression based approach to explore the causal link between PAs and poverty outcomes. We model poverty metrics in 2011 while controlling for baseline poverty in 2001 to avoid constructing models that can generate spurious correlations (Brett, 2004). This approach yields similar parameter estimates for our treatment variables as those generated when modelling absolute change (S2.2). The pre-processing of data using matching methods optimizes the balance of covariates across treated and control units, and is useful when imbalance between treatment and control is an issue for traditional causal inference techniques (Ho *et al.* 2007). We used genetic matching with replacement, which performs well when covariates have skewed distributions (Diamond and Sekhon 2013).

We performed all of our statistical analyses in *R* version 3.3.2 (2013) using the “Matchit” package (Ho *et al.* 2007). We used post-matching standardized mean differences of <0.25 as an acceptable balance between treatment and control groups for each covariate (Stuart 2010, see Figures S2.3-S2.5). We then performed an Ordinary Least Squares (OLS) regression to adjust for remaining imbalances in covariate distributions (Ho *et al.* 2007). When modelling extreme poverty, we implemented a two-step hurdle model (Cragg 1971) using matched binomial regressions to first model the incidence of extreme poverty, and then OLS regressions to model the magnitude of extreme poverty in those VDCs in which extreme poverty occurs. We first measured the average impact of our treatments (protection) on our response variables (poverty, extreme poverty and inequality in 2011). We then subset and separately matched PAs in each tourism intensity category (high, intermediate or low) to assess the impact of tourism intensity. PAs with high tourism levels were all designated before 2001, so we only performed this subgroup analysis on PAs established before 2001. We conducted robustness checks to test for spillover effects from unprotected VDCs adjacent to a PA (defined as the treatment) into unprotected control VDCs that are not adjacent to a PA (S2.4), and spatial autocorrelation (S2.5); results are robust to spillover and spatial autocorrelation unless stated otherwise.

2.3.3 Heterogeneity analysis

We assessed if PA impacts were moderated by travel time to the nearest tourism hub (PA entrance, major and minor trekking route) and elevation, which affects livelihood choices (greater range of options in the lowlands, including commercial agriculture) and tourism options (safaris in the lowlands, trekking in the mountains). We used partial linear modelling (PLM - Yatchew 1998; Hanauer 2015) to assess heterogeneous impacts along the gradients of our moderating factors following methods described in Ferraro *et al.* (2011) and Hanauer and Canavire-Bacarreza (2015). In a first step, we controlled for confounding factors using a linear regression. In the second stage, we employed a nonparametric locally weighted scatter plot smoothing (LOESS) to estimate the non-parametric relationship between moderator and outcome. This method allows us to estimate the impact of PAs on our outcome variables as a function of our moderator variables of interest (elevation and travel times to the nearest PA entrance, major and minor trekking routes) while holding other covariates constant.

2.4 Results

i) Average impact on poverty, extreme poverty and inequality

We found no evidence that PAs exacerbated poverty in Nepal. In fact, matched protected VDCs (defined using the 10% threshold and established before 2001) had significantly lower poverty in 2011 than unprotected VDCs (coef. = -0.03, S.E. = 0.02, P = 0.027; Figure 2.2A). Poverty was not exacerbated when raising the protection threshold to 70% (coef. = -0.06, S.E. = 0.03, P = 0.060; Figure 2.2A). For PAs established after 2001 we found no evidence of positive or negative impacts of PAs on overall poverty (Figure 2.3A). Models without matching showed similar patterns (Table S2.7).

PAs established before and after 2001 reduced the incidence of extreme poverty. For PAs established before 2001, this result was significant for our 10% protection threshold (coef. = -0.95, S.E. = 0.38, *P* = 0.012; Figure 2.2B) and was accentuated by raising the threshold to 70% (coef. = -3.51, S.E. = 1.20, *P* = 0.003; Figure 2.2B). For PAs established after 2001, this result was not significant using a 10% protection threshold, but was significant after raising the protection threshold to 70% (coef. = -2.82, S.E. = 1.18, *P* = 0.018; Figure 2.2B). We found no significant impact of protection on the magnitude of extreme poverty (Table S2.8). Results from models without matching showed the same patterns for PAs established before and after 2001 (Table S2.7).

We found no consistent evidence that inequality was influenced by PAs established before or after 2001, using either 10% or 70% protection thresholds (Figure 2.2C). Models without matching indicate that PAs established before 2001 reduced inequality, while PAs established after 2001 increased inequality (Table S2.7), but these difference were not significant after controlling for spatial autocorrelation (Table S2.6).

ii) Tourism intensity

PAs with high tourism levels significantly reduced overall poverty (coef. = -0.05, S.E. = 0.02, *P* = 0.023; Figure 2.2A), while PAs with low tourism levels had no significant effect on poverty. However, PAs with low tourism levels significantly alleviated extreme poverty (coef. = -2.80, S.E. = 1.12, *P* = 0.013; Figure 2.2B) and decreased inequality (coef. = -1.01, S.E. = 0.43, *P* = 0.023; Figure 2.2C).

iii) Heterogeneity – travel time to PA entrance and trekking route

Travel time to a PA entrance had no impact on poverty (Figure S2.7) and inequality (Figure S2.6), while reductions in the incidence of extreme poverty were greater closer to a PA entrance (Figure 2.3D). Travel time to a minor and major trekking route moderated the influence of PAs on poverty, with significant reductions only occurring in VDCs close to the trekking route (Figures 2.3A and 2.3B). Incidence of extreme poverty was lower further away from a minor trekking routes (Figure 2.3D), but was not influenced by travel time to a major trekking route (Figure S2.7). Inequality was not influenced by proximity to major or minor trekking routes (Figure S2.6).

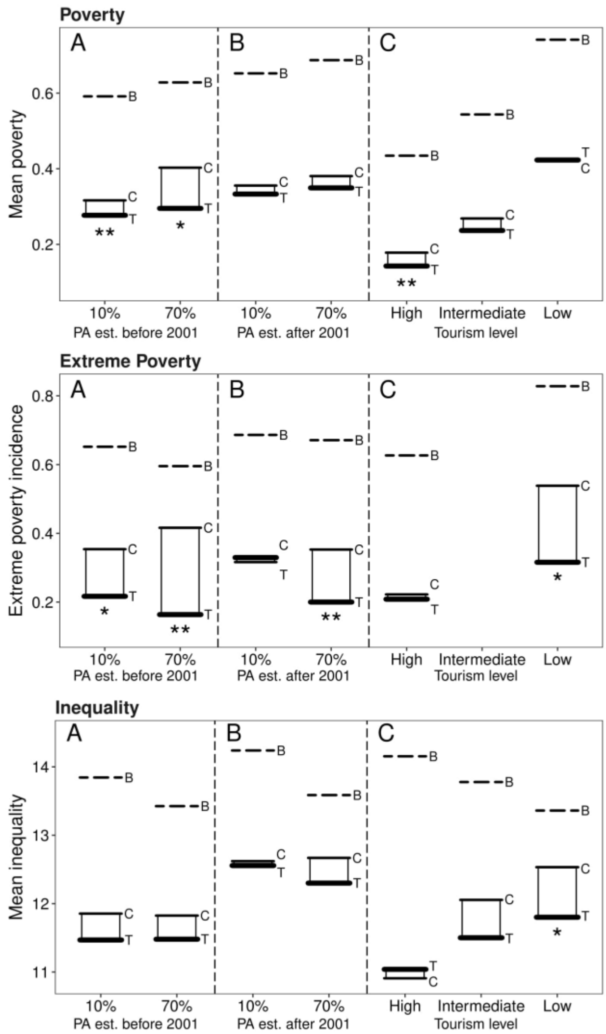

Figure 2.2. Estimated impacts of protected areas (PAs) on poverty, extreme poverty and inequality in Village Development Committees (VDCs) in Nepal for PAs established before 2001 (A), PAs established between 2001 and 2011 (B), and according to level of tourism (C). Poverty, extreme poverty and inequality measurements are based on a multidimensional poverty index. Dashed lines (B) represent mean baseline (2001) of VDCs, thick lines (T) represent treatment i.e. PAs, thin lines (C) represent counterfactual controls without protection. Significance: **** P < 0.001, *** P < 0.01, ** P < 0.05, * <0.1.

iv) Heterogeneity – elevation

Our PLM results do not show significant heterogeneous impacts of PAs on extreme poverty and inequality as a function of elevation (Figures 2.3F and S2.6). PAs established before 2001 reduced poverty to a greater extent at low elevations than high elevations (Figure 2.3E).

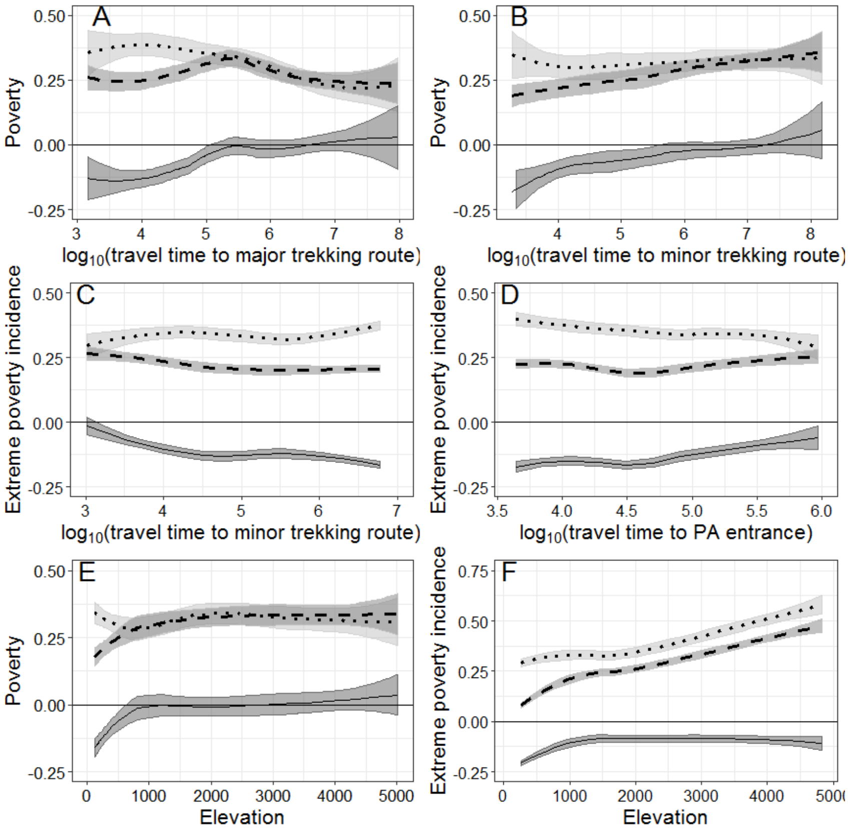

Figure 2.3. Partial linear models: Impact of protected areas (PA) on poverty (A-C) and extreme poverty (D-F) in Village Development Committees (VDCs) in Nepal for PAs established before 2001, conditional on travel time to major trekking route (A), minor trekking route (B-C) and PA entrance (D), and PA impacts conditional on elevation (E-F). Poverty measurements are based on a multidimensional poverty index. Dashed lines represent protected VDCs, dotted lines counterfactual controls without protection, and solid lines the difference between treatment and counterfactual estimates (negative values indicate reductions in poverty).

2.5 Discussion

Nepali PAs typically reduced poverty, concurring with previous research elsewhere (Andam *et al.* 2010; Hanauer and Canavire-Bacarreza 2015). Crucially, PAs reduced extreme poverty without deepening inequalities. This finding is particularly important as creating pathways out of extreme poverty is more difficult than tackling less extreme poverty (Halder and Mosley 2004). Our findings suggest that PAs are able to provide pathways out of extreme poverty in remote areas, challenging previous evidence that PA policies only benefit community elites (Agrawal and Gupta 2005).

PAs with high tourism levels reduced poverty without exacerbating extreme poverty and inequality, while PAs with low tourism levels reduced extreme poverty and inequality but had no impact on overall poverty. These results suggest that the poorest receive the greatest benefits from small-scale tourism, contrasting with previous suggestions that tourism increases inequalities (West *et al.* 2006). We provide further evidence for beneficial impacts from tourism by showing that poverty reductions in PAs only occurred close to trekking routes. This suggests that redistribution policies (that 30-50% of PA revenue is spent on local community development; Heinen and Shrestha 2006) may not fully address spatial biases in which communities benefit from tourism in PAs. Notably, however, the impact of PAs on reducing extreme poverty increased with distance from minor trekking routes that are typically located in remote areas with little development potential that can benefit from park redistribution policies. Future studies should specifically assess if, where and how these policies influence PA poverty outcomes.

Distance from PA entrances had no impact on extreme poverty inside PAs, but increased extreme poverty outside PAs. This suggests localised negative spillovers, with PA residents living close to PA entrances receiving benefits that people living equally close to entrances outside of the PA miss out on. Other research on PA spillover effects show similar patterns of heterogeneity (Robalino *et al.*, 2017; Pfaff and Robalino, 2017), with tourism benefits only occurring close to PA entrances (Robalino and Villalobos 2015). Indeed, our analyses indicate that benefits of protection do not spread to neighbouring unprotected VDCs. Redistribution policies might thus need to target communities inside and outside protected area more equally.

Time since establishment moderated the effect of PAs on our measures of poverty. PAs established after 2001 did not show the same significant social benefits as PAs established before 2001, although in newer PAs we observe a trend towards lower extreme poverty and inequality. This pattern is expected if there are time lag effects that arise because communities need to adjust to new regulations imposed by PAs and the new opportunities provided by them, and for the tourism industry to develop. The reduced benefits of more recently established PAs are unlikely to be associated with changes in management regimes as these have been constant across all Nepali PAs since the 1990s (Bhattarai et al., 2017). Notably, an increase in the threshold used to define a protected VDC (from 10% to 70%) accentuated our main findings. This suggests that communities in VDCs that have restrictions placed on activities across a larger proportion of their land do not experience adverse impacts on poverty metrics, thus larger protected areas may deliver greater economic benefits. Finally, impacts of PAs were similar across a wide range of elevations indicating that PAs can deliver socio-economic impacts even in areas that typically support livelihoods that are less compatible with nature conservation, such as agriculture.

Our study makes a number of important contributions. First, we demonstrate not only that PAs in Nepal reduce poverty and extreme poverty, but that they do so without increasing inequality. These benefits occur even in lowland regions with high capacity for alternative land-uses, and when capacity for alternative livelihoods is reduced by protecting larger proportions of land. Second, we find that tourism is a key driver of PA benefits, but that reductions of extreme poverty are possible even in marginalised areas with limited tourism potential. Finally, we find no evidence that socio-economic benefits of PAs spread to people living outside, but close to, PAs. Addressing this by adjusting PA's revenue redistribution policies, could increase the benefits for these communities and reduce conflict between local communities and PA's conservation objectives (Oldekop *et al.* 2016). Nepal's PA management policy to promote social welfare via redistribution of PA revenues, gained through tourism and other activities, is similar to policies in other countries including Thailand (Sims 2000) and Kenya (Walpole and Leader-Williams, 2001) suggesting that our findings may also apply elsewhere.

2.6 Acknowledgements

BdB was funded by the Grantham Centre for Sustainable Futures and a scholarship from the Green Economy Network award from the University of Sheffield. J.A.O was supported by a European Union funded Marie Curie Outgoing Fellowship (FORCONEPAL)

2.7 References

Agrawal, A. & Gupta, K. (2005). Decentralization and participation: The governance of common pool resources in Nepal's Terai. *World Dev.*, 33, 1101–1114.

Alkire, S. & Foster, J. (2011). Counting and multidimensional poverty measurement. *J. Public Econ.*, 95, 476–487.

Andam, K.S., Ferraro, P.J., Sims, K.R.E., Healy, A. & Holland, M.B. (2010). Protected areas reduced poverty in Costa Rica and Thailand. *Proc. Natl. Acad. Sci. U. S. A.*, 107, 9996–10001.

Ban, N.C., Mills, M., Tam, J., Hicks, C.C., Klain, S., Stoeckl, N., Bottrill, M.C., Levine, J., Pressey, R.L., Satterfield, T. & Chan, K.M.A. (2013). A social-ecological approach to conservation planning: embedding social considerations. *Front. Ecol. Environ.*, 11, 194–202.

Bhattarai, B.R., Wright, W., Poudel, B.S., Aryal, A., Yadav, B.P. & Wagle, R. (2017). Shifting paradigms for Nepal's protected areas: history, challenges and relationships. *J. Mt. Sci.*, 14, 964–979.

Brett, M.T. (2004). When is a correlation between non-independent variables "spurious"? *Oikos*, 105, 647–656.

Brockington, D. & Wilkie, D. (2015). Protected areas and poverty. *Philos. Trans. R. Soc. B-Biological Sci.*, 370.

Cragg, J.G. (1971). Some statistical models for limited dependent variables with application to demand for durable goods. *Econometrica*, 39, 829–844.

de Sherbinin, A. (2008). Is poverty more acute near parks? An assessment of infant mortality rates around protected areas in developing countries. *Oryx*, 42, 26–35.

Diamond, A. & Sekhon, J.S. (2013). Genetic matching for estimating causal effects: A general multivariate matching method for achieving balance in observational studies. *Rev. Econ. Stat.*, 95, 932–945.

Ferraro, P.J. & Hanauer, M.M. (2015). Through what mechanisms do protected areas affect environmental and social outcomes? *Philos. Trans. R. Soc. B-Biological Sci.*, 370.

Ferraro, P.J., Hanauer, M.M. & Sims, K.R.E. (2011). Conditions associated with protected area success in conservation and poverty reduction. *Proc. Natl. Acad. Sci. U. S. A.*, 108, 13913–13918.

Fisher, B. & Christopher, T. (2007). Poverty and biodiversity: measuring the overlap of human poverty and the biodiversity hotspots. *Ecol. Econ.*, 62, 93–101.

Gentle, P. & Maraseni, T.N. (2012). Climate change, poverty and livelihoods: adaptation practices by rural mountain communities in Nepal. *Environ. Sci. Policy*, 21, 24–34.

Halder, S.R. & Mosley, P. (2004). Working with the ultra-poor: learning from BRAC experiences. *J. Int. Dev.*, 16, 387–406.

Hanauer, M.M. & Canavire-Bacarreza, G. (2015). Implications of heterogeneous impacts of protected areas on deforestation and poverty. *Philos. Trans. R. Soc. B-Biological Sci.*, 370.

Hanauer, M.M. (2015). Partial Linear Differencing Model [WWW Document]. URL http://www.sonoma.edu/users/h/hanauer/Rfunctions/rfunctions.html

Heinen, J.T. & Shrestha, S.K. (2006). Evolving policies for conservation: an historical profile of the protected area system of Nepal. *J. Environ. Plan. Manag.*, 49, 41–58.

Ho, D.E., Imai, K., King, G. & Stuart, E.A. (2007). Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. *Polit. Anal.*, 15, 199–236.

IUCN & UNEP-WCMC. (2016). The World Database on Protected Areas (WDPA).

Joppa, L.N. & Pfaff, A. (2009). High and Far: Biases in the Location of Protected Areas. *PLoS One*, 4.

Lokshin, M. & Ravallion, M. (2000). Welfare Impacts of the 1998 Financial Crisis in Russia and the Response of the Public Safety Net. *Econ. Transit.*, 8, 269–295.

Ministry of Culture T.C.A. (2013). *Nepal Tourism Statistics*. Singha Durbar, Kathmandu.

Miranda, J.J., Corral, L., Blackman, A., Asner, G. & Lima, E. (2014). *Effects of Protected Areas on Forest Cover Change and Local Communities: Evidence from the Peruvian Amazon*.

Oldekop, J.A., Holmes, G., Harris, W.E. & Evans, K.L. (2016). A global assessment of the social and conservation outcomes of protected areas. *Conserv. Biol.*, 30, 133–141.

Pfaff, A. & Robalino, J. (2017). Spillovers from Conservation Programs. *Annu. Rev. Resour. Econ.*, 9, 299–315.

Robalino, J., Pfaff, A. & Villalobos, L. (2017). Heterogeneous Local Spillovers from Protected Areas in Costa Rica. *J. Assoc. Environ. Resour. Econ.*, 4, 795–820.

Robalino, J. & Villalobos, L. (2015). Protected areas and economic welfare: an impact evaluation of national parks on local workers' wages in Costa Rica. *Environ. Dev. Econ.*, 20, 283–310.

Sims, K.R.E. (2010). Conservation and development: Evidence from Thai protected areas. *J. Environ. Econ. Manage.*, 60, 94–114.

Sims, K.R.E. & Alix-Garcia, J.M. (2017). Parks versus PES: Evaluating direct and incentive-based land conservation in Mexico. *J. Environ. Econ. Manage.*, 86, 8–28.

Spiteri, A. & Nepal, S.K. (2008). Evaluating local benefits from conservation in Nepal's Annapurna Conservation Area. *Environ. Manage.*, 42, 391–401.

Stuart, E.A. (2010). Matching Methods for Causal Inference: A Review and a Look Forward. *Stat. Sci.*, 25, 1–21.

Team R.C. (2013) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Walpole, M.J. & Leader-Williams, N. (2001). Masai Mara tourism reveals partnership benefits. *Nature*, 413, 771.

West, P., Igoe, J. & Brockington, D. (2006). Parks and peoples: The social impact of protected areas. *Annu. Rev. Anthropol.*, 35, 251–277.

Yatchew, A. (1998). Nonparametric regression techniques in economics. *J. Econ. Lit.*, 36, 669–721.

Yergeau, M.-E., Boccanfuso, D. & Goyette, J. (2017). Linking conservation and welfare: A theoretical model with application to Nepal. *J. Environ. Econ. Manage.*, 85, 95–109.

2.7 Supplementary information

S1 Poverty Metrics

Poverty is considered a multidimensional concept that includes more dimensions than the traditionally used measures of household consumption and income (Alkire and Santos, 2014). We used the Nepali national census data (2001 – 520,624 households surveyed; 2011 – 841,467 households surveyed) to construct multi-dimensional poverty (MDP) indices based on indicators of household health, education and living standards (Table S2.1). We derived three MDP indices: poverty, (MDP>0.33), extreme poverty (MDP>0.66) and inequality (standard deviation of household MDPs). The MDP poverty index is closely correlated with the $1.25/day poverty line (Alkire and Santos, 2014), suggesting that it is also representative of household income and consumption. Other studies have recommended the GINI-index as a measure of inequality between rich and poor (Halpern et al. 2013), but this is based on income and ignores non-monetary based forms of assets and well-being that are crucial for more holistic assessment of poverty. We therefore used the standard deviation of household MDPs as a measure of inequality.

Table S2.1. The dimensions, indicators, deprivation cut-offs and weights of the multidimensional poverty index

Dimension	Indicator	Deprived if:	Weight
Health (weight 0.33)	1. Child mortality	The household experienced death of one or more children aged 5 years or younger in the last 12 months	0.166
	2. Premature mortality	The household experienced premature death (death below the period life expectancy) in the last 12 months	0.166
Education (weight 0.33)	1. School absence	The household has a school-aged child not going to school (a school-aged child is defined as aged between six and sixteen years)	0.166
	2. Years of schooling	The household has a member of eleven years or older that has not completed five years of schooling	0.166
Living standards (weight 0.33)	1. Cooking fuel	The household uses wood or dung as cooking fuel	0.0833
	2. Electricity	The household has no access to electricity	0.0833
	3. Source of water	The household has no access to clean water	0.0833
	4. Improved Sanitation	The household does not have access to improved sanitation (flush toilet or similar)	0.0833

S2 Model definition

Instead of modelling poverty in 2011 as the response variable it would also be possible to estimate absolute change in poverty (poverty 2011-poverty 2001) as the response variable. However, as we show below, this model is functionally the same, and does not affect the estimates of the treatment effect.

Under our approach our model to measure poverty in 2011 (*Yi*) takes the following form:

$Y_t \sim \beta_0 + \beta_1 * Y_{t-1} + \beta_2 * X_t + \epsilon$. (equation 1)

Here we model Yi as a function of baseline poverty Yi-1 and treatment Xi. β1 represents the coefficient of baseline poverty, β2 represents the treatment effect, and ϵ is a normally distributed error term. Using absolute change as the response is equivalent to subtracting baseline poverty from both sites to generate the following model:

$Yt - Y_{t-1} \sim \beta_0 + \beta_1 * Y_{t-1} - 1 * Y_{t-1} + \beta_2 * X_t + \epsilon$. (equation 2)

Which simplifies to:

$Yt - Y_{t-1} \sim \beta_0 + (\beta_1 - 1) * Y_{t-1} + \beta_2 * X_t + \epsilon$. (equation 3)

Such a model that measures absolute change would have different parameter estimates for baseline poverty (β1-1) and could generate spurious correlations (Brett, 2004), but the parameter coefficient of the focal treatment effect (β2) remains unchanged following this reformulation of the model (see Table S2.2 for comparison).

Table S2.2. Comparison of the association between treatment and poverty metrics when modelling poverty using alternative response variables, i.e. i) absolute poverty in 2011 (equation 1) and ii) modelling change in poverty (equation 2).

	Model using poverty metric 2011 as response variable Coefficient ± Standard error (*P*-value)	**Model using absolute change as response variable** Coefficient ± Standard error (*P*-value)
Poverty		
PAs established before 2001 (10% overlap, *n = 214*)	-0.035 ± 0.015 (*P* = 0.026)	-0.035 ± 0.015 (*P* = 0.026)
PAs established after 2001 (10% overlap, *n=150*)	-0.006 ± 0.017 (*P* = 0.733)	-0.006 ± 0.017 (*P* = 0.733)
Inequality		
PAs established before 2001 (10% overlap ,*n = 223*)	-0.164 ± 0.226 (*P* = 0.467)	-0.164 ± 0.226 (*P* = 0.467)
PAs established after 2001 (10% overlap ,*n = 149*)	0.172 ± 0.259 (*P* = 0.508)	0.172 ± 0.259 (*P* = 0.508)
Extreme poverty incidence		
PAs established before 2001 (10% overlap ,*n = 227*)	-0.951 ± 0.381 (*P* = 0.012)	-0.951 ± 0.381 (*P* = 0.012)
PAs established after 2001 (10% overlap ,*n = 153*)	-0.102 ± 0.413 (*P* = 0.805)	-0.102 ± 0.413 (*P* = 0.805)
Extreme poverty magnitude		
PAs established before 2001 (10% overlap ,*n = 227*)	-0.015 ± 0.027 (*P* = 0.577)	-0.015 ± 0.027 (*P* = 0.577)
PAs established after 2001 (10% overlap ,*n = 153*)	-0.000 ± 0.002 (*P* = 0.954)	-0.000 ± 0.002 (*P* = 0.954)

S3 Confounding factors and transformations

Our selection of covariates follows other research on protected areas (Andam et al. 2010; Canavire-Bacarreza and Hanauer 2013). Table S2.3 presents the confounding factors used in our analysis, and the rationale for including them. PA entrances and trekking route locations were determined from guidebooks (Armington 2001; McConnachie et al. 2012) and the Forestry Nepal website (www.forestrynepal.org). We generated travel time estimates by adapting the approach used to calculate travel time to major cities developed by the Joint Research Centre (JRC) (Nelson 2008). To do so, we combined the Nepali Survey Departments road data, the global land cover dataset developed by JRC (Bartholome and Belward 2005) and the ASTER global digital elevation model version 2 (Tachikawa et al. 2011) to produce a travel time raster for population centres of different sizes and each of our tourism surrogate indicators (PA entrance and major/minor trekking route). These were then multiplied by 1 km resolution population estimates (LandScan, Oak Ridge National Laboratory; Bright and Coleman 2001) and divided by the total population in each VDC in order to take into account the non-random distribution of population centres, especially in mountainous areas.

We tested for colinearity among variables to avoid redundancy in the data. We removed travel time to population centres larger than 25,000 inhabitants from the analysis because it was highly correlated with travel time to population centres higher than 10,000 ($r = 0.75$). Slope and elevation were highly colinear ($r = 0.79$) across the entire dataset but there was much spatial variation in the strength of this relationship so both were retained as confounding factors. We $\log_{10}$-transformed covariates with highly skewed distributions to reduce kurtosis and associated problems of outliers disproportionately influencing the analysis.

Table S2.3. Confounding factors used in analyses. The unit of analysis is the Village Development Committee (VDC).

Confounding factor	Definition and data source	Rationale for inclusion
Elevation and slope	Mean elevation (m.) and slope (degree), calculated from ASTER DEM version 2 (Tachikawa et al. 2011)	Slope and elevation influence the distributions of PAs and people's livelihood decisions (Gentle and Maraseni 2012)
Area*	Size of the VDC (km^2). Data: Nepali Department of Home Affairs	Size of the administrative area has been found to correlate with poverty outcomes (Andam et al. 2008)
Forest cover*	Proportion of forested area in VDC in 2000 with more than 10% (UNFAO). Data: High-resolution forest cover change dataset version 1.0 (Hansen et al. 2013)	Forest cover influences PA allocation and welfare as many people depend on forests for their livelihoods (Angelsen et al. 2014)
Precipitation	Mean levels of precipitation in individual VDCs using the 30 arc-seconds resolution WorldClim current precipitation dataset (Hijmans et al. 2005)	Precipitation affects forest dynamics, agricultural activity and related land use choices (Duncan et al. 2013)
Population density*	Population size for each VDC at baseline 2001 as recorded in the Nepali census	Population density affects both welfare and PA allocation (Wittemyer et al. 2008)
Agricultural effort	Mean number of months that household members above school age (>16 years of age) dedicate to agriculture in VDC (Nepal census)	Agriculture is a leading cause of land-cover change and an important source of livelihoods for local communities (Angelsen 2010)
International migration*	Proportion of households within each VDC with at least one or more household members above school age (>16 years) living abroad (Nepal census)	International migration influences socioeconomic patterns through remittances (Barham and Boucher 1998)
Community forest management size	Proportion of VDC under community forest management (Nepali Department of forests)	Community forest management is an important factor influencing socioeconomic outcomes (Acharya 2002; Persha et al. 2011)
Community forest management age	Mean number of years since community forest management was established (Nepali Department of forests)	
Travel time to pop. centre: - 1,000 inhabitants* - 10,000 inhabitants * - 25,000 inhabitants † - 50,000 inhabitants * - District HQ*	Travel time calculated using the method described above for tourism surrogates (Nepali Survey Department)	Access to markets and services (e.g., technical assistance, health centres) can significantly influence people's livelihoods, PA establishment decisions, and other related land-use choices (Angelsen 2010)
District	Exact match based on the district which VDCs are located in	Districts have significant autonomy in decision-making and planning, and the majority of donor funded development projects are targeted at the district-level

* $\log_{10}$ transformed
† Dropped due to collinearity

S4 Spillover analysis

We performed a spillover analysis because households in unprotected VDCs that are close to protected VDCs might also be positively or negatively influenced by PAs (Pfaff and Robalino, 2017). To test whether spillover effects occurred we first selected all unprotected VDCs (<1% overlap with a PA) adjacent to a protected VDCs (i.e. over 10% of their area overlapping a PA). We defined these adjacent VDCs as treated and matched them to other control VDCs that were not adjacent to a protected VDC. Under the null hypothesis of zero spillover there should be no treatment effect in the matched sample. We repeated these steps for all three social outcome variables, and for PAs established before and after 2001. Our spillover analyses do not detect any significant spillover effects from protected VDCs into adjacent unprotected VDCs (Table S2.4).

Table S2.4. Regression results for spillover analysis of protected areas (PAs). All covariates were used in the analysis but their coefficients are not shown here

	Coefficient	Standard error	*P*
Poverty			
PAs established before 2001	0.013	0.018	0.458
PAs established after 2001	0.031	0.022	0.166
Inequality			
PAs established before 2001	-0.053	0.282	0.850
PAs established after 2001	0.284	0.331	0.394
Extreme poverty incidence			
PAs established before 2001	-0.010	0.594	0.987
PAs established after 2001	1.377	0.968	0.155
Extreme poverty magnitude			
PAs established before 2001	0.006	0.008	0.456
PAs established after 2001	-0.004	0.004	0.302

S5 Spatial autocorrelation

Spatial autocorrelation can distort parameter estimates. Spatial autocorrelation is measured by defining a weight matrix based on a measure of distance between observations. We calculated the weight matrix based on inverted distance between VDCs to ensure that nearby VDCs are allocated more weight than VDCs further away. We performed all spatial autocorrelation analyses in *R* version 3.3.2 (2013) using the "spdep" package (Bivand 2015). For our poverty and inequality outcome variables we calculated the Global Moran's *I* index in the model residuals to detect spatial autocorrelation in our data. In the case of significant spatial autocorrelation we then determined the type of dependence (spatial lag or spatial error) based on the Lagrange multiplier diagnostics. If necessary, we then corrected for spatial autocorrelation based on the Lagrange multiplier diagnostics (Anselin 2002; Dormann et al. 2007) using a spatial lag or spatial error model. Spatial lag occurs when observations are not independent as the dependent variable in a location is influenced by neighbouring independent variables. Spatial error indicates that the error terms across different spatial units are correlated and could implicate that omitted spatial variables might affect inference (Anselin, 2002). In the case of the extreme poverty incidence response variable spatial eigenvector mapping (Griffith and Peres-Neto, 2006) indicated negligible spatial autocorrelation in our models. We estimated spatial autocorrelation for all models with and without matching. However, the weights that are generated by the matching interfere with the weights that are used in the spatial autocorrelation models. In this case we used the matched subsample without the matching weights and applied the spatial models to this reduced dataset.

Overall, spatial analyses confirmed our main findings. Testing for spatial autocorrelation in our matched models for PAs established before 2001 showed that spatial lag was present in our data (P <0.05; Table S2.5). However, after applying a spatial lag model the results confirm our main finding that PAs before 2001 reduced poverty and did not exacerbate inequality. We did not find evidence of spatial autocorrelation in our matched models for PAs established after 2001 and also found no significant spatial eigenvectors in our models for extreme poverty incidence, indicating that spatial autocorrelation did not distort our parameter estimations.

Our models without matching showed significant spatial error in all models (Table S2.6). Correcting using a spatial error model did not remove all spatial error, although Moran's I values were greatly reduced. We found that PA impacts on inequality were heavily influenced by spatial error. After employing a spatial error model we found no impact of PAs on inequality, in line with our models with matching. Applying a spatial error model did not result in significant differences in poverty and extreme poverty impacts for both PAs established before and after 2001.

Table S2.5. Regression results of spatial autocorrelation analysis – models with matching. The coefficient standard error and P value of treatment (i.e. protection) and global Moran's I values are shown for the post-matched regression, and the spatial lag model when the global Moran's *I* value is significant in the post-matched regression. All covariates were used in the analysis, but their coefficients are not shown here.

		Coefficient	Standard error	*P*	Global Moran's *I* (*P*)
Poverty					
PAs established before 2001 (*n = 214*)	Post-matched regression	-0.034	0.015	0.027	0.036 (0.036)
	Spatial lag model	-0.041	0.015	0.005	-0.071 (0.959)
PAs established after 2001 (*n=150*)	Post-matched regression	-0.006	0.017	0.733	-0.076 (0.588)
Inequality					
PAs established before 2001 (*n = 223*)	Post-matched regression	-0.164	0.226	0.467	0.154 (<0.001)
	Spatial error model	-0.265	0.225	0.238	0.001 (0.449)
PAs established after 2001 (*n = 149*)	Post-matched regression	0.172	0.259	0.508	-0.049 (0.886)
Extreme poverty incidence					
PAs established before 2001 (*n = 227*)	Post-matched regression	-0.951	0.381	0.012	-0.043 (0.868)
PAs established after 2001 (*n = 153*)	Post-matched regression	-0.102	0.413	0.805	-0.013 (0.599)
Extreme poverty magnitude					
PAs established before 2001 (*n = 64*)	Post-matched regression	-0.015	0.027	0.577	-0.053 (0.414)
PAs established after 2001 (*n = 51*)	Post-matched regression	0.000	0.002	0.954	-0.154 (0.875)

Table S2.6. Regression results of spatial autocorrelation analysis – models without matching. The coefficient standard error and P value of treatment (i.e. protection) and global Moran's I values are shown for the post-matched regression, and the spatial lag model when the global Moran's *I* value is significant in the post-matched regression. All covariates were used in the analysis, but their coefficients are not shown here.

		Coefficient	Standard error	*P*	Global Moran's *I* (*P*)
Poverty					
PAs established before 2001	Regression	-0.039	0.009	<0.001	0.134 (<0.001)
(*n = 3714*)	Spatial error model	-0.022	0.009	0.029	0.017 (<0.001)
PAs established after 2001	Regression	-0.020	0.011	0.072	0.132 (<0.001)
(*n = 3628*)	Spatial error model	-0.015	0.012	0.201	0.015 (<0.001)
Inequality					
PAs established before 2001	Regression	-0.554	0.142	<0.001	0.159 (<0.001)
(*n = 3714*)	Spatial error model	-0.224	0.149	0.133	0.013 (<0.001)
PAs established after 2001	Regression	0.336	0.172	0.051	0.152 (<0.001)
(*n = 3628*)	Spatial error model	0.088	0.181	0.626	0.012 (<0.001)
Extreme poverty incidence					
PAs established before 2001	Regression	-0.955	0.232	<0.001	0.001 (0.006)
(*n = 3714*)	Spatial eigenvector mapping	-1.024	0.241	<0.001	0.001 (0.001)
PAs established after 2001	Regression	-0.512	0.252	0.043	0.007 (0.023)
(*n = 3628*)	Spatial eigenvector mapping	-0.463	0.263	0.078	0.002 (0.236)
Extreme poverty magnitude					
PAs established before 2001	Regression	0.000	0.003	0.906	0.010 (<0.001)
(*n = 1153*)	Spatial error model	0.001	0.003	0.707	-0.020 (0.997)
PAs established after 2001	Regression	-0.003	0.003	0.306	0.076 (<0.001)
(*n = 1149*)	Spatial error model	-0.002	0.003	0.500	-0.027 (0.999)

Table S2.7. The coefficient, standard error and P value of treatment (protection) in regression models of poverty, inequality and extreme poverty without matching of covariates (hence the larger sample size). Extreme poverty results are presented as a hurdle model: 1. a binomial model to estimate extreme poverty incidence, and 2. a linear regression model of the magnitude of extreme poverty in those VDCs where extreme poverty is present (only for 10% overlap due to sample sizes). All matching confounding factors were included in the regressions, but their coefficients are not shown here.

	Coefficient	**Standard error**	***P***
Poverty			
PAs established before 2001 ≥ 10% overlap (*n = 3714*)	-0.039	0.009	<0.001
PAs established before 2001 ≥ 70% overlap (*n = 3632*)	-0.043	0.012	<0.001
PAs established after 2001 ≥ 10% overlap (*n = 3628*)	-0.020	0.011	0.072
PAs established after 2001 ≥ 70 % overlap (*n = 3589*)	-0.027	0.014	0.054
Inequality			
PAs established before 2001 ≥ 10% overlap (*n = 3714*)	-0.554	0.142	<0.001
PAs established before 2001 ≥ 70% overlap (*n = 3632*)	-0.757	0.191	<0.001
PAs established after 2001 ≥ 10% overlap (*n = 3628*)	0.336	0.172	0.051
PAs established after 2001 ≥ 70% overlap (*n = 3589*)	0.383	0.213	0.073
Extreme poverty incidence and magnitude			
PAs established before 2001 ≥ 10% overlap 1. Extreme poverty incidence (binomial model) (*n = 3714*)	-0.955	0.232	<0.001
PAs established before 2001 ≥ 10% overlap 2. Extreme poverty magnitude (linear regression) (*n = 1153*)	0.000	0.003	0.906
PAs established before 2001 ≥ 70% overlap 1. Extreme poverty incidence (binomial model) (*n = 3632*)	-1.123	0.320	<0.001
PAs established after 2001 ≥ 10% overlap 1. Extreme poverty incidence (binomial model) (*n = 3628*)	-0.512	0.252	0.043
PAs established after 2001 ≥ 10% overlap 2. Extreme poverty magnitude (linear regression) (*n = 1149*)	-0.003	0.003	0.305
PAs established after 2001 ≥ 70% 1. Extreme poverty incidence (binomial model) (*n = 3589*)	0.984	0.351	0.005

Table S2.8. The coefficient, standard error and P value of treatment (protection) in matched regression models of the magnitude of extreme poverty (i.e. second stage of the hurdle model). All covariates were used in the analysis, but their coefficients are not shown here.

	Coefficient	Standard error	*P*
Extreme poverty magnitude			
PAs established before 2001 ≥ 10% overlap Linear regression (*n = 64*)	-0.015	0.027	0.577
PAs established after 2001 ≥ 10% overlap Linear regression (*n = 51*)	0.000	0.002	0.954

Table S2.9. Trekking routes included in our analysis. Descriptions from Armington (2008). Classifications are based on descriptions that indicate popularity.

#	Name	Description	Classification
1	Annapurna Circuit Trek	"there are teashops and lodges every couple of hours"	Major
2	Annapurna Panorama Trek	"this is a popular trip"	Major
3	Annapurna Sanctuary Trek	"you will rarely walk long without finding refreshment or accommodation"	Major
4	Everest Base Camp Trek	"tens of thousands of trekkers storm the trail every year"	Major
5	Gosainkund Trek	"popular alternative"	Major
6	Langtang Valley Trek	"one of Nepal's big three treks"	Major
7	Short treks near Pokhara	"most popular region among trekkers"	Major
8	Gokyo Trek	"less crowded"	Minor
9	Helambu Trek	"few crowds"	Minor
10	Khopra Ridge Trek	"trek away from the popular trails"	Minor
11	Mustang Trek	“now over 3000 trekkers a year visit"	Minor
12	Around Manaslu Trek	"without the crowds"	Minor
13	Beni to Dolpo	"more blue sheep than people"	Minor
14	Ganja La Trek	"the most challenging and dangerous"	Minor
15	Jumla to Dunai	"facilities are very sparse"	Minor
16	Kagmara La Trek	"little information"	Minor
17	Kanchenjunga Treks	"remote Alpine country", "facilities are limited"	Minor
18	Limi Valley trek	"little-visited corner of Nepal"	Minor
19	Lukla to Tumlingtar	"little-used route"	Minor
20	Makalu Base Camp Trek	"rugged and remote"	Minor
21	Nar-Phu	"far removed in both time and space"	Minor
22	Phoksumdo Lake to Shey Gompa	"truly adventurous"	Minor
23	Phoksumdo Lake Treks	"wild, high country without rescue or communication facilities"	Minor
24	Rara Lake Trek	"well off the beaten track"	Minor
25	Shivalaya to Lukla Trek	"only 1000 trekkers walk this route a year"	Minor
26	Tamang Heritage Trail	"off the main tourist route"	Minor
27	Three Passes Trek	"only for the truly adventurous"	Minor

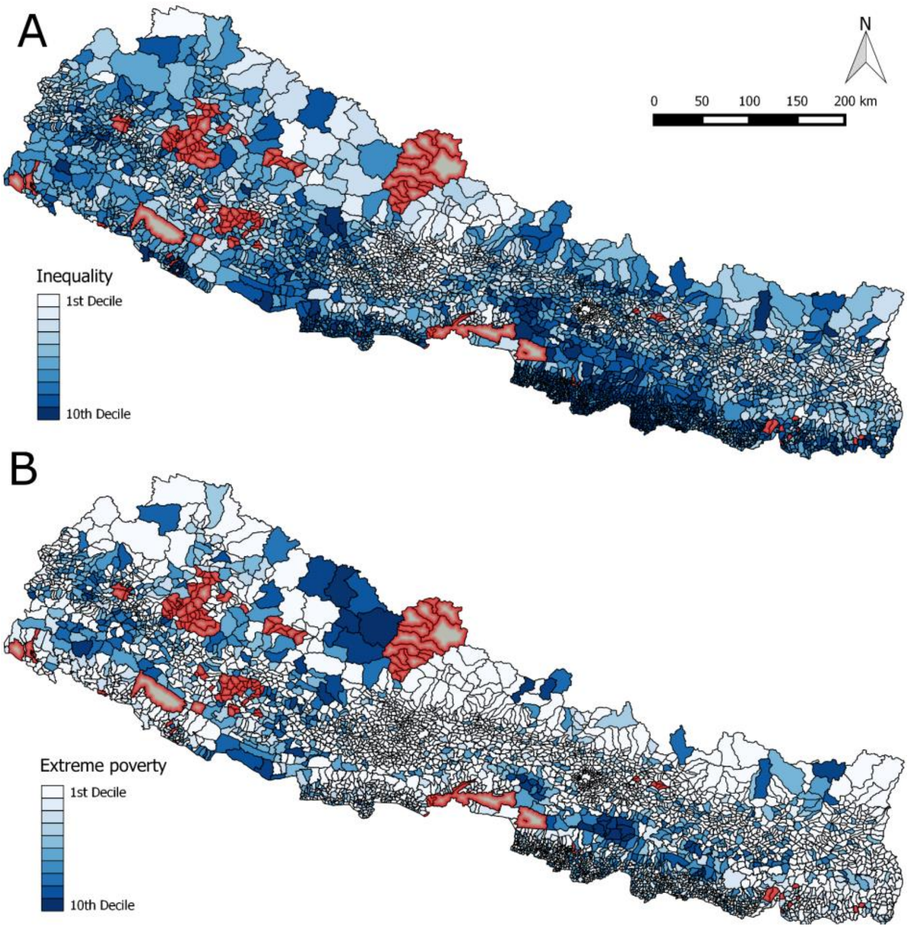

Figure S2.1. Inequality (A) and extreme poverty (B) in 2011 in Nepal, based on multidimensional poverty index (MPI). Data are presented as deciles. Grey areas with red contours represent excluded Village Development Committees (reasons for exclusion include missing data due to armed conflict and instances of inconsistent data from the Nepali Department of Forests concerning community forestry data).

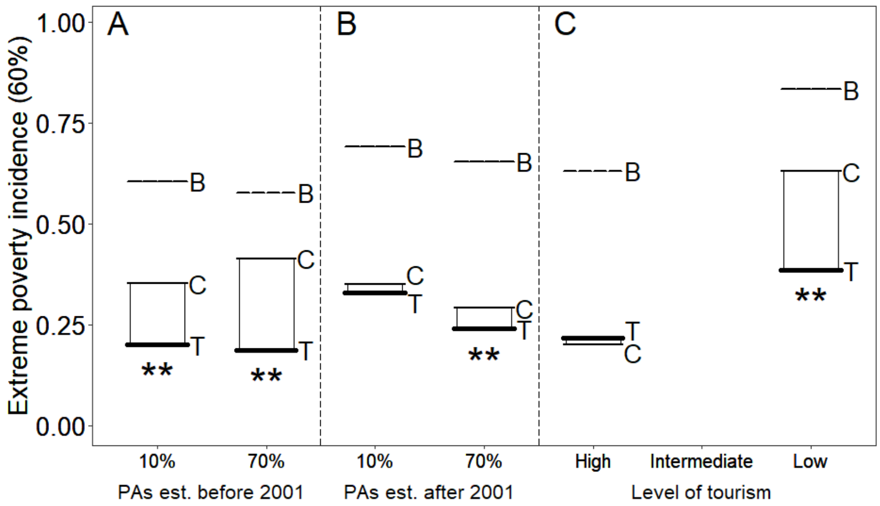

Figure S2.2. Estimated impacts of protected areas (PAs) on extreme poverty in Nepali Village Development Committees (VDCs) using the alternative 60 % threshold for PAs established before 2001 (A), PAs established between 2001 and 2011 (B), and according to level of tourism (C). Dashed lines (B) represent mean baseline (2001) of VDCs, thick lines (T) represent treatment i.e. PAs, thin lines (C) represent counterfactual controls without protection. Significance: **** *P* < 0.001, *** *P* < 0.01, ** *P* < 0.05, * *P* <0.1.

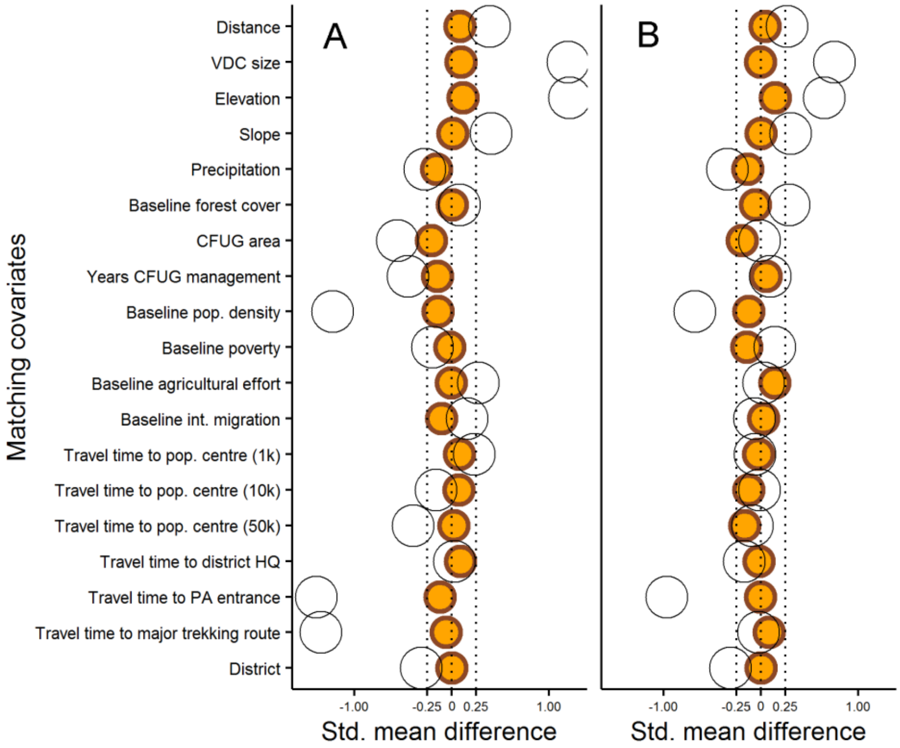

Figure S2.3. Balance of covariates before and after genetic matching, using protection as treatment and poverty as outcome variable. (A) PAs established before 2001. (B) PAs established after 2001. Standardized mean difference for confounding covariates before (open circles) and after matching (orange circles). Dotted lines indicate acceptable standard mean difference after matching to reach balance (Stuart 2010).

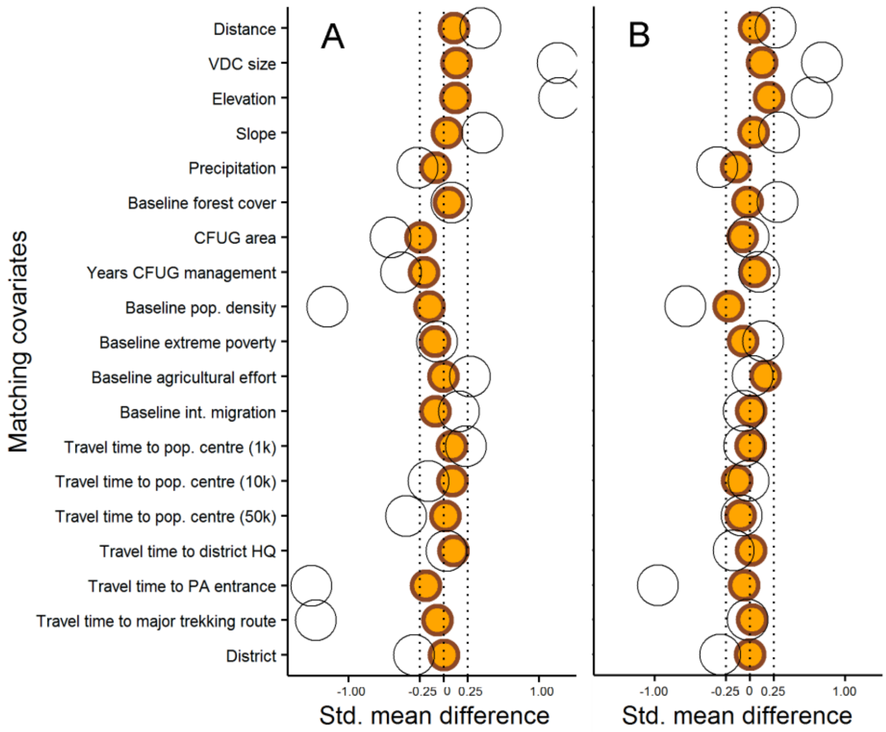

Figure S2.4. Balance of covariates before and after genetic matching, using protection as treatment and extreme poverty as outcome variable. (A) PAs established before 2001. (B) PAs established after 2001. Standardized mean difference for confounding covariates before (open circles) and after matching (orange circles). Dotted lines indicate acceptable standard mean difference after matching to reach balance (Stuart 2010).

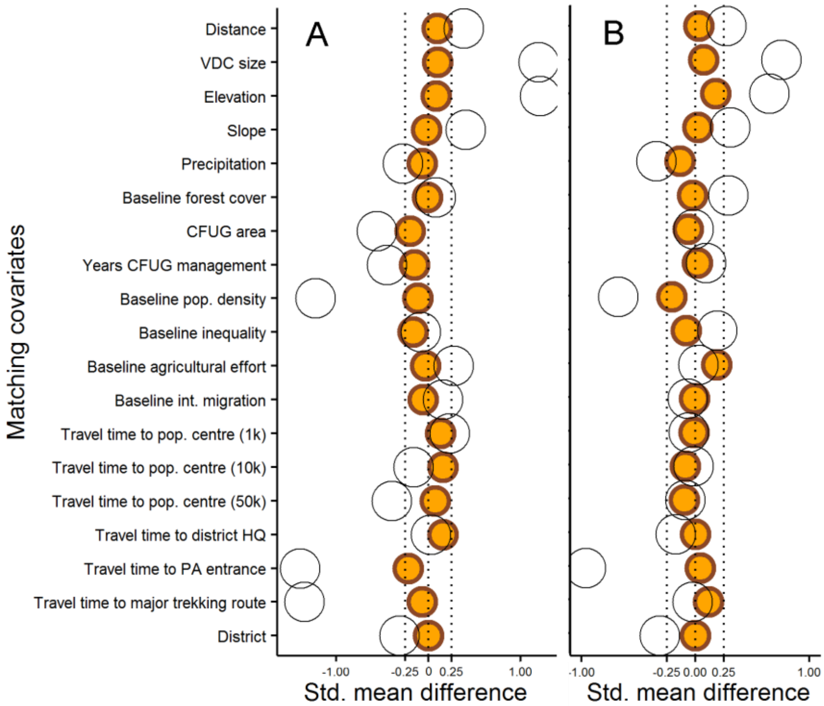

Figure S2.5. Balance of covariates before and after genetic matching, using protection as treatment and inequality as outcome variable. (A) PAs established before 2001. (B) PAs established after 2001. Standardized mean difference for confounding covariates before (open circles) and after matching (orange circles). Dotted lines indicate acceptable standard mean difference after matching to reach balance (Stuart 2010).

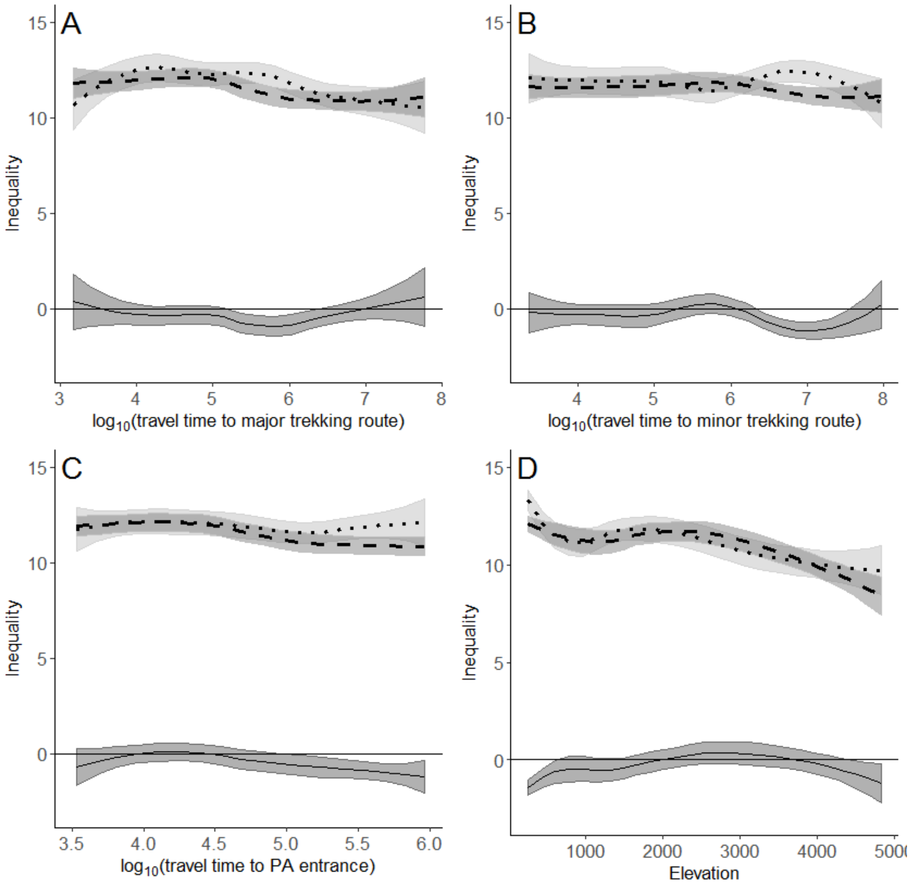

Figure S2.6. Partial linear model: Impact of protected areas (PA) on inequality in Village Development Committees (VDCs) in Nepal for PAs established before 2001, conditional on travel time to minor trekking route (A), major trekking route (B), PA entrance (C) and elevation (D). Poverty measurements are based on a multidimensional poverty index. The dashed line represents protected VDCs, the dotted line counterfactual control without protection, and the solid line the difference between protected and counterfactual (negative values indicate PAs reduced inequality).

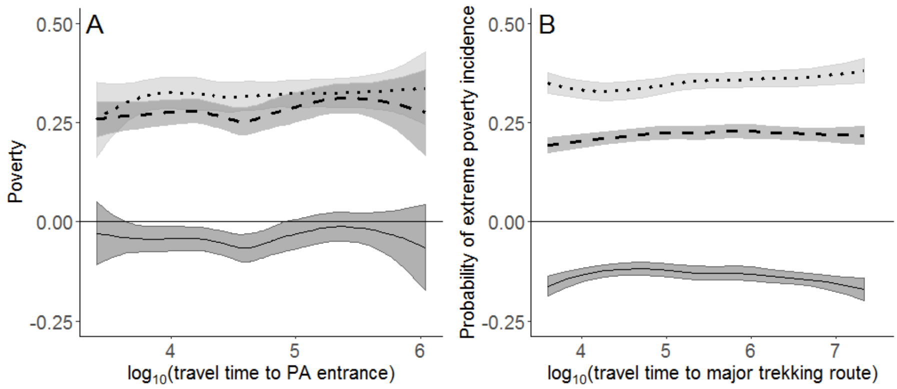

Figure S2.7. Partial linear model: Impact of protected areas (PA) on poverty (A) and extreme poverty (B) in Village Development Committees (VDCs) in Nepal for PAs established before 2001, conditional on travel time to PA entrance (A) and major trekking route (B). Poverty measurements are based on a multidimensional poverty index. The dashed line represents protected VDCs, the dotted line counterfactual control without protection, and the solid line the difference between protected and counterfactual (negative values indicate PAs reduced poverty/extreme poverty).

References

Acharya, K.P. (2002). Twenty-four years of community forestry in Nepal. *Int. For. Rev.*, 4, 149–156.

Andam, K.S., Ferraro, P.J., Pfaff, A., Sanchez-Azofeifa, G.A. & Robalino, J.A. (2008). Measuring the effectiveness of protected area networks in reducing deforestation. *Proc. Natl. Acad. Sci. U. S. A.*, 105, 16089–16094.

Andam, K.S., Ferraro, P.J., Sims, K.R.E., Healy, A. & Holland, M.B. (2010). Protected areas reduced poverty in Costa Rica and Thailand. *Proc. Natl. Acad. Sci. U. S. A.*, 107, 9996–10001.

Angelsen, A. (2010). Policies for reduced deforestation and their impact on agricultural production. *Proc. Natl. Acad. Sci. U. S. A.*, 107, 19639–19644.

Angelsen, A., Jagger, P., Babigumira, R., Belcher, B., Hogarth, N.J., Bauch, S., Boerner, J., Smith-Hall, C. & Wunder, S. (2014). Environmental Income and Rural Livelihoods: A Global-Comparative Analysis. *World Dev.*, 64, S12–S28.

Anselin, L. (2002). Under the hood - Issues in the specification and interpretation of spatial regression models. *Agric. Econ.*, 27, 247–267.

Armington, S. (2001). *Trekking in the Nepal Himalaya*. Lonely planet.

Barham, B. & Boucher, S. (1998). Migration, remittances, and inequality: estimating the net effects of migration on income distribution. *J. Dev. Econ.*, 55, 307–331.

Bartholome, E. & Belward, A.S. (2005). GLC2000: a new approach to global land cover mapping from Earth observation data. *Int. J. Remote Sens.*, 26, 1959–1977.

Bivand, R. (2015). Spatial Dependence: Weighting Schemes, Statistics and Models R package spdep. *Compr. R Arch. Netw.*, version 0.

Bright, E.A. & Coleman, P.R. (2001). LandScan 2000. LandScan.

Canavire-Bacarreza, G. & Hanauer, M.M. (2013). Estimating the Impacts of Bolivia's Protected Areas on Poverty. *World Dev.*, 41, 265–285.

Dormann, C.F., McPherson, J.M., Araujo, M.B., Bivand, R., Bolliger, J., Carl, G., Davies, R.G., Hirzel, A., Jetz, W., Kissling, W.D., Kuehn, I., Ohlemueller, R., Peres-Neto, P.R., Reineking, B., Schroeder, B., Schurr, F.M. & Wilson, R. (2007). Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. *Ecography (Cop.).*, 30, 609–628.

Duncan, J.M.A., Biggs, E.M., Dash, J. & Atkinson, P.M. (2013). Spatio-temporal trends in precipitation and their implications for water resources management in climate-sensitive Nepal. *Appl. Geogr.*, 43, 138–146.

Gentle, P. & Maraseni, T.N. (2012). Climate change, poverty and livelihoods: adaptation practices by rural mountain communities in Nepal. *Environ. Sci. Policy*, 21, 24–34.

Griffith, D. & Peres-Neto, P. (2006). The Flexibility of Eigenfunction Spatial Analyses. *Ecology*, 87, 2603–2613.

Halpern, B.S., Klein, C.J., Brown, C.J., Beger, M., Grantham, H.S., Mangubhai, S., Ruckelshaus, M., Tulloch, V.J., Watts, M., White, C. & Possingham, H.P. (2013). Achieving the triple bottom line in the face of inherent trade-offs among social equity, economic return, and conservation. *Proc. Natl. Acad. Sci. U. S. A.*, 110, 6229–6234.

Hansen, M.C., Potapov, P. V, Moore, R., Hancher, M., Turubanova, S.A., Tyukavina, A., Thau, D., Stehman, S. V, Goetz, S.J., Loveland, T.R., Kommareddy, A., Egorov, A., Chini, L., Justice, C.O. & Townshend, J.R.G. (2013). High-Resolution Global Maps of 21st-Century Forest Cover Change. *Science (80-.).*, 342, 850–853.

McConnachie, J., Meghji, S. & Reed, D. (2012). *The rough guide to Nepal*. Rough Guides UK.

Nelson A. (2008) Estimated travel time to the nearest city of 50,000 or more people in year 2000. Global Environment Monitoring Unit-Joint Research Centre of the European Commission, Ispra Italy, http://forobs.jrc.ec.europa.eu/products/gam/.

Persha, L., Agrawal, A. & Chhatre, A. (2011). Social and Ecological Synergy: Local Rulemaking, Forest Livelihoods, and Biodiversity Conservation. *Science (80-.).*, 331, 1606–1608.

Pfaff, A. & Robalino, J. (2017). Spillovers from Conservation Programs. *Annu. Rev. Resour. Econ.*, 9, 299–315.

Stuart, E.A. (2010). Matching Methods for Causal Inference: A Review and a Look Forward. *Stat. Sci.*, 25, 1–21.

Tachikawa, T., Kaku, M., Iwasaki, A., Gesch, D.B., Oimoen, M.J., Zhang, Z., Danielson, J.J., Krieger, T., Curtis, B. & Haase, J. (2011). *ASTER global digital elevation model version 2-summary of validation results*.

Wittemyer, G., Elsen, P., Bean, W.T., Burton, A.C.O. & Brashares, J.S. (2008). Accelerated human population growth at protected area edges. *Science (80-.).*, 321, 123–126.

Chapter 3:

Differences in Deforestation and Poverty between Sustainable Use Protected Areas, Agricultural Landholdings and Presence of Mines in the Brazilian Amazon

Bowy den Braber, Javier Godar, Christoph Nolte, Karl L. Evans, Johan A. Oldekop

B.d.B., J.A.O. and K.L.E. designed research; B.d.B. performed research and analysed data; J.G. and C.N. provided data; B.d.B. wrote the paper with input from all authors

Met opmerkingen [b10]: Author contribution

3.1 Abstract

Understanding how land use decisions drive deforestation and poverty is key for policy makers, but studies that assess trade-offs between multiple alternatives remain rare. We compare effects of sustainable use protected areas (SU PAs) on deforestation and poverty in the Brazilian Amazon with two competing land uses: agricultural landholdings and mining. On average, SU PAs were effective in reducing deforestation without exacerbating poverty, but this masks important heterogeneity. There were no differences in deforestation between SU PAs and smallholders or remote areas, but SU PAs were much more effective in comparison to large landholdings than average comparisons. Reduced deforestation outcomes did not lead to poverty exacerbation by SU PAs, even in comparison to areas dominated by large landholdings. SU PAs also reduced deforestation in comparison to mining sites, but at the expense of locally reduced incomes and higher inequality. Mining effects on income and deforestation spread beyond the present mine.

3.2 Introduction

Deforestation and poverty are at the core of many sustainability challenges, including climate change, biodiversity loss and food security (Foley et al. 2005; Naeem et al. 2016). Gaining a deeper understanding of how policies can drive transitions to sustainable land use practices is essential to address increasing land scarcity and competition among alternative land uses (Ferraro et al. 2018a). Despite substantial progress in measuring the consequences of land-use policies and interventions, evaluations that move beyond binary comparisons of two alternative land-uses and acknowledge that there are multiple alternatives remain rare (but see e.g. Heilmayr and Lambin (2016)). Elucidating comparative effects of a range of alternative land uses is essential to capture the range of options available to policy makers, and generate better theoretical and empirical understanding of environmental and socio-economic trade-offs of land-use alternatives within socio-ecological systems.

We quantify deforestation and poverty outcomes of three common land-use options within tropical forested landscapes, using the Brazilian Legal Amazon (BLA) as a case study. Specifically, we contrast outcomes arising when land is managed as sustainable use protected areas (SU PAs) – a rapidly growing conservation and development strategy (Poutzols et al. 2014), agriculture and mining. Our analysis further distinguishes between agricultural land dominated by smallholders and large landholders, as these represent very divergent forms of agriculture and the relative contributions of these groups to deforestation and socio-economic outcomes is frequently debated. Large-scale agricultural production is responsible for most deforestation and produce typically flows to distant consumers with little contribution to local poverty reduction (Meyfroidt 2018). Conversely, the contribution of smallholders to deforestation is lower, and while financial benefits are more localised (Pokorny et al. 2013), revenues are smaller as smallholders are often disconnected from agricultural markets (Barrett et al. 2001). Mining is a significant driver of tropical deforestation but estimates of its impacts are highly variable (Edwards et al. 2014; Austin et al. 2019), and its local effects on poverty are highly contested (Bebbington et al. 2008; von der Goltz & Barnwal 2018). For example, mining in Peru increased wages, but also increased inequality (Loayza & Rigolini 2016).

Met opmerkingen [b11]: Example to clarify point made

Our analysis focuses on deforestation and poverty outcomes for the period 2000-2010. Despite previous work on BLA PAs (e.g. Nolte et al. 2013; Soares-Filho, 2010; Kere et al. 2017) we are unaware of studies that have assessed both the environmental and socioeconomic effects of PAs relative to the main competing land uses in the region. A small number of studies from other countries have compared PA impacts with those of competing land uses but focus exclusively on deforestation outcomes (Schleicher 2017; Panlasigui et al. 2018). Some studies have also assessed PA effects on poverty but only in comparison to alternative forest-based interventions, such as payments for ecosystem services (e.g. Sims and Alix-Garcia 2017), and have ignored alternative uses that have greater potential to contribute to deforestation whilst providing key resources for people and traditional economic activity, such as agricultural products or mined resources.

Comparing effects of PAs to alternative land uses is particularly important in the Brazilian Amazon whose forests have very high biodiversity value and provide the world's largest terrestrial carbon stock (Nepstad et al. 2009). Socio-economic benefits of PAs in other regions appear to be largely driven by eco-tourism (e.g. Naidoo et al. 2019; den Braber et al. 2018), but such benefits are likely to be less prevalent in the Brazilian Amazon as the majority remains largely inaccessible to visitors. Moreover, due to deforestation Brazil has become the world's largest emitter of forest-based greenhouse gases (Tyukavina et al. 2017). Deforestation initially peaked in the early 2000s as a result of expansion of soy and cattle production lands, influenced by technological advancements and favourable market conditions (Nepstad et al. 2014a; Richards et al. 2014). Deforestation rates dropped after 2004 when the Brazilian government launched the Plan to Combat Deforestation in the Amazon (PPCDAM in Portuguese acronym) that included the expansion of PAs among other interventions (Arima et al. 2014). More recently, the Brazilian government has implement a range of initiatives that have been linked to increasing deforestation. Land-use within the Brazilian Amazon was historically decided directly by central government (Hecht 1985). This role of central government continues today and includes rolling back environmental protection (Pack et al. 2016), major development projects including infrastructure projects (Ferreira et al. 2014) and agrarian settlement schemes aimed at smallholder farms (Pacheco 2009). These settlement schemes are a source of deforestation (Schneider & Peres 2015), but large landholders are responsible for the largest part of the deforestation even though the relative contribution of smallholders has increased (Godar et al., 2014). Mining is the second largest cause of Amazonian deforestation after farming related activities, contributing about 10% of recent deforestation (Sonter et al. 2017). Gold mining in the Amazon has been linked to mercury exposure, but broader socio-economic impacts of mining are not well studied (Malm 1998). Rapidly shifting and unstable land tenure designations have caused disputes and conflicts over who owns Brazilian lands (Sparovek et al. 2019). The relationship between environmental conservation and other actors in Brazil has been tense for decades, highlighted by the increase of PA downgrading, downsizing and degazettement (PADDD) events in the Amazon (Pack et al. 2016) and the potential for mining within PA boundaries (Ferreira et al. 2014). This trend is predicted to accelerate following recent government pledges to loosen regulations and further open the Amazon to agriculture and mining (Escobar 2019).

We compiled a high-spatial resolution, longitudinal dataset that combined data on land-use, deforestation and poverty. We generated deforestation estimates from the well-established PRODES Brazilian Amazon deforestation dataset (INPE 2017) and the global forest loss dataset from the University of Maryland (Hansen et al. 2013), and estimated four well-established poverty indicators (income, inequality, literacy and sanitation) for 5,545 census tracts from population censuses. These census tracts are the smallest spatial unit for which socio-economic data are available within the BLA and comprise the entire BLA. We compiled property sizes of landholders from agricultural census data and approved mining concessions from SIGMINE (DNPM 2012) to define different subsets of comparison groups. We combined these with spatial data on SU PAs that allow controlled access and extraction of resources. We focus on SU PAs because they allow controlled extraction of resources, and thus provide a more comparable alternative scenario to agriculture and mining than strictly protected areas. In addition, they comprise a major component of protected land in the BLA (36%) and are the most rapidly growing type of PA globally and in the BLA in recent years (Soares-Filho et al.

2010; IUCN & UNEP-WCMC 2016). We use well-established matching and regression techniques to control for potential biophysical, socio-economic and political confounders that could influence our outcomes (S3.1). Our analysis focuses on the expansion of SU PAs established after 2000 as this enables us to assess changes in our key outcome variables following establishment of PAs. We also use different definitions of treatment (from 10% (187 CTs) to 50% (112 CTs) of the CT under protection) to test whether PA effects are moderated by the amount of protected land.

3.3 Results

Using a quasi-experimental approach our analysis across 5,545 census tracts demonstrate that the expansion of SU PAs in the Brazilian Amazon in comparison to unprotected land, *on average*, significantly reduced deforestation between 2000 and 2010 ($P < 0.001$; Fig. 1) without exacerbating poverty. This equates to roughly 1,750 hectares on average per census tract, attributed to SU PAs established after 2000. We found no significant impact of SU PAs established after 2000 on any of our poverty indicators (income ($P = 0.749$), inequality ($P = 0.487$), literacy ($P = 0.392$) and sanitation ($P = 0.318$) Fig.1b). These broad patterns remained for both deforestation and poverty indicators, with similar effect sizes, when raising the definition of protected census tracts from 10% to 50% overlap with SU PAs (Figure S3.1A-S3.2). The global forest loss dataset from the University of Maryland also showed this pattern, but estimates of PA impact were higher (Figure S3.1C).

Crucially, taking heterogeneity in land use in the control census tracts into account reveals significant context dependency in the effect of SU PAs on both deforestation and poverty outcomes. We used agricultural census data to define agricultural controls. CTs with less than 10% of their land settled by landholders were classified as “remote”. We split the remaining CTs according to landholding size, and classified the census tracts based on the most frequent type of property size. We used four categories of property size: 1) “very small” properties smaller than 10 ha, 2) “small” properties between 10 and 50 ha, 3) “medium” properties between 50 and 200 ha, and 4) “large” properties bigger than 200 ha. We find that SU PAs established after 2000 have not reduced deforestation in comparison to remote areas ($P = 0.178$) and areas dominated by landholders with very small property sizes, i.e. < 10 ha ($P = 0.293$), but were successful in reducing deforestation in comparison to areas dominated by landholders with property sizes > 10 ha (Figure 3.1). Effect sizes of avoided deforestation per census unit differed in comparison to small (landholdings of 10 to 50 ha; ~2,500 ha of avoided deforestation), medium (landholdings of 50 to 200 ha; ~6,600 ha of avoided deforestation) and large (landholdings larger than 200 ha; ~4,200 ha of avoided deforestation) landholders. Importantly, these effects of SU PAs are greater than those detected in the initial analysis that pooled all unprotected land in the control treatment (1.4 times greater for small, 3.7 times for medium and 2.4 times for large landholdings). When using the global forest loss dataset, rather than the PRODES dataset, we find a similar pattern of effects, but do see that SU PAs had limited (<100ha) but significant effect on reducing deforestation in comparison to remote areas (Fig. S3.1C). The differential effect of PAs when taking land holding size of the control units into account remains similar when raising the threshold of the amount of land protected to 50% (Figure S3.1A). We tested whether these results were influenced by differences in the amount of land that is settled by agricultural landholders, but we found no significant differences in the proportion of settled land between the categories (F = 0.69, df = 121, *P* = 0.56, Figure S3.4A).

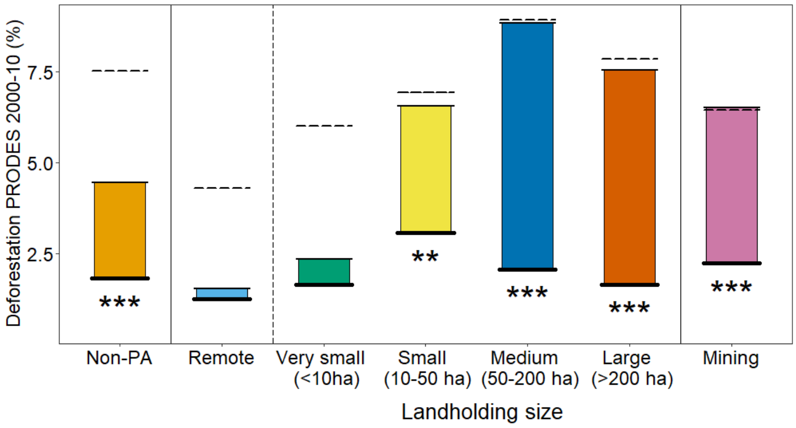

Figure 3.1. Estimated impacts of sustainable use (SU) protected areas established after 2000 on deforestation in comparison to non-protection, agricultural controls split by dominant size of landholdings, and registered mining sites, in census tracts (CTs) in the Brazilian Amazon. Non-PA controls include all the CTs with less than 1% of their land protected. Other categories related to non-protected controls with specific land use pattern. Remote areas are those with less than 10% of the CT settled. Dominant agricultural landholder size is determined by the category that contains the most properties per CT. Mining sites are defined as CTs with recorded licensed mining activities present after 2000. Thick lines represent treatment i.e. PAs, thin lines represent modelled values in controls matched to the treatment across all ten covariates. Dashed lines represent average level in non-matched control

group. Significant differences between treatment and matched controls: *** $P < 0.001$, ** $P < 0.01$.

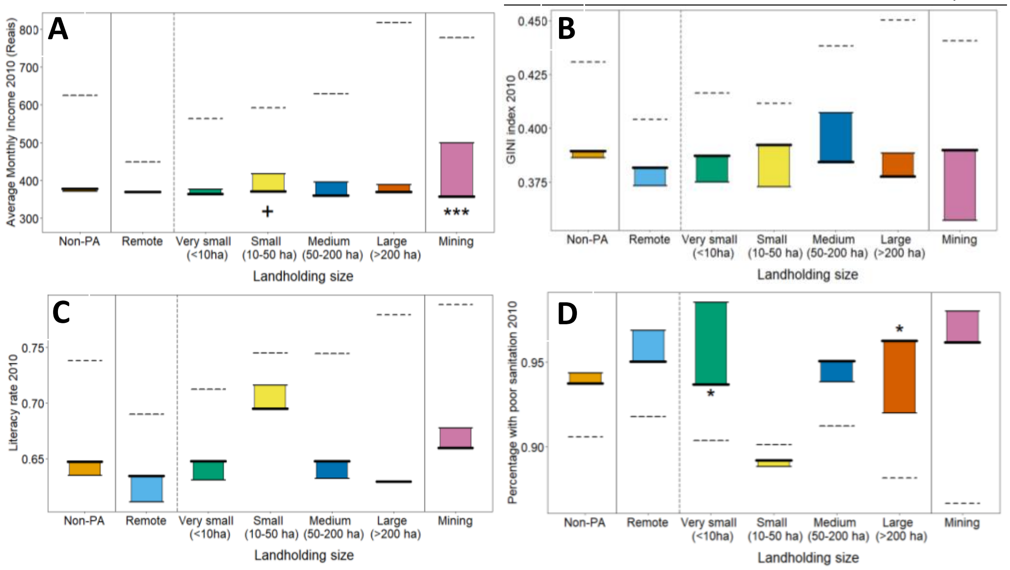

Figure 3.2. Estimated impacts of sustainable use protected areas (SU PAs) established after 2000 on income (A), inequality (B), literacy (C), and sanitation (D), in comparison to non-protection, agricultural controls split by dominant size of landholdings, and registered mining sites, in census tracts (CTs) of the Brazilian Amazon. Non-PA controls include all the CTs with less than 1% of their land protected. Other categories related to non-protected controls with specific land use pattern. Remote areas are those with less than 10% of the CT settled. Dominant agricultural landholder size is determined by the category that contains the most properties per CT. Mining sites are defined as CTs with recorded licensed mining activities present after 2000. Thick lines represent treatment i.e. SU PAs, thin lines represent modelled values in controls matched to the treatment across all ten covariates. Dashed lines represent average level in non-matched control group. Significance: *** $P < 0.001$, * $P < 0.05$, + $P < 0.1$.

Agriculture in the BLA is dominated by pasture (i.e. livestock) or growing arable crops (such as soy); both types of agriculture are important drivers of deforestation although their relative contributions have varied over time with pasture being more important during our focal time period (Nepstad et al. 2014a). We find that deforestation rates are similar in SU PAs and census tracts dominated by small agricultural land holdings, and there is generally slightly, but significantly, less pasture in these census tracts (Figure S3.4B). However, we find that deforestation rates remain similar in SU PAs and census tracts dominated by small agricultural land holdings when these are restricted to those dominated by pasture (Figure S3.1B).

We found no consistent evidence that areas dominated by agriculture differed from SU PAs established after 2000 in their ability to alleviate poverty (Figure 3.2). There was marginal, non-significant, evidence that SU PAs reduced income compared to areas dominated by small landholders (P = 0.098; Figure 3.2A), but we found no differences in effects on literacy and inequality of SU PAs and other size of agricultural landholding (Figure 3.2C). Effects of SU PAs on sanitation relative to agricultural landholdings had consistently small effect sizes and were very variable in the direction of effects (Figure 3.2D).

When we compare SU PAs to CTs with an active legal mining concession we find trade-offs between deforestation (Figure 3.1) and poverty (Figure 3.2) outcomes. SU PAs established after 2000 were successful in reducing forest loss in comparison to census tracts with mining activity (~3500ha; $P < 0.001$; Figure 3.1), but mining sites are associated with higher income (P = 0.021; Figure 3.2A) and lower inequality (P = 0.028; Figure 3.2B). We find no difference in the effects of mining and SU PAs on literacy or sanitation (Figures 3.2C-D). These patterns were confirmed after raising the threshold to 50% land protected (Figure S3.2). Our main analyses defined mined census tracts as those in which mines were located, but mining effects on deforestation can extend up to 75km from the mining site (Sonter et al. 2017). We thus repeated analyses when changing the definition of mined census tracts as those in which a mine was located to any census tract within 75 km, 50 km, 25 km, 10 km and 5 km from a mine. These analyses find that SU PAs reduce deforestation relative to mining across all definitions of mined census tracts up to and including those within 50km of a mine, but not those within 75km of a mine (Fig. 3.3A). Mining, relative to SU PAs, also increased income when mined census tracts were defined as those within 5 and 10 km of a mine (Fig. 3.3B). Literacy rates were higher in census tracts 5-25km away from a mining sites, but we found no significant difference at other distances (Fig. S3.5). We found no significant effect of mines on inequality and sanitation under different definitions of mined census tracts (Fig. S3.5).

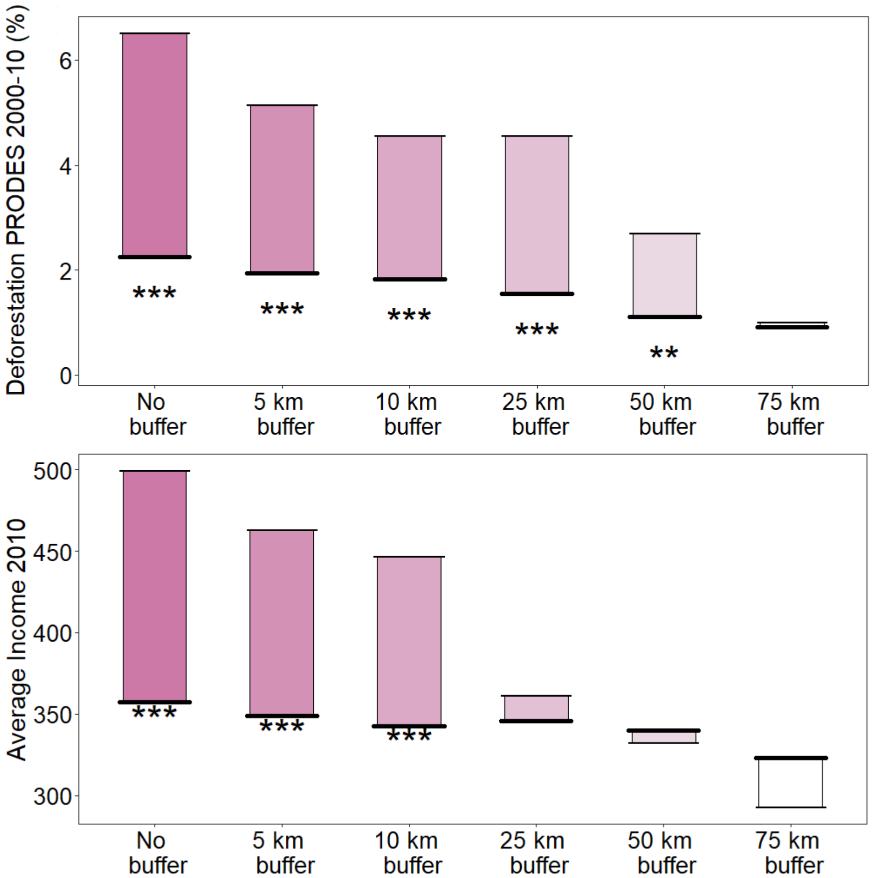

Figure 3.3. Estimated impacts of sustainable use (SU) protected areas established after 2000 in comparison to mining concessions across different buffer distances, in census tracts of the Brazilian Amazon. Thick lines represent treatment i.e. PAs, thin lines represent modelled values in controls matched to the treatment across all ten covariates. Significance: *** $P < 0.001$, * $P < 0.05$.

3.4 Discussion

Our study demonstrates that focusing on PA effects without considering the nature of land-use in unprotected controls masks considerable variation in in the consequences for conservation (deforestation) and socioeconomic outcomes (poverty). Such heterogeneity in land-use is, however, typically ignored in studies of PA impacts which prevents policy makers assessing if alternative land use decisions would generate better or worse outcomes that protection across multiple and diverse sustainability objectives. For example, even though previous studies have shown that PAs can reduce poverty (e.g. Andam et al. 2010), alternative land uses could be preferable, depending on development priorities, if they were able to reduce poverty even more.

A key finding is that deforestation outcomes of SU PAs do not significantly differ when compared to areas dominated by very small land holdings (<10 ha) and remote areas with little settled land (although agriculture in such areas can still lead to the loss of disturbance sensitive species; Socolar et al. 2019). Previous research suggests that large landholders are the largest contributors to deforestation in the BLA, although the share of deforestation caused by smallholders has increased recently (Godar et al. 2014; Richards & VanWey 2016) and we demonstrate that SU PAs can still reduce deforestation relative to agricultural areas dominated by land holdings between 10 and 50 ha. The lack of differences in deforestation between SU PAs and remote areas contrasts with the disproportionate location of PAs in remote areas (Joppa & Pfaff 2009), confirming previous suggestions that PAs have previously been located in areas where it is politically convenient to do so rather than where they are most effective (Rodrigues et al. 2004). Notably, however, the SU PAs that were established during the study period are located less remotely than already established PAs (Nolte et al. 2013). Moreover, projected future increases in the accessibility of remote areas due to improved transport infrastructure (Barber et al. 2014) suggest that such PAs will, however, be well placed to play a key role in reducing deforestation in the near future.

The environmental benefits arising from reduced deforestation outcomes of SU PAs are unlikely to trade-off against the capacity for poverty alleviation as we find no evidence of poverty exacerbation by SU PAs, even in comparison to areas dominated by larger landholdings. This suggests that economic benefits from agricultural activities in these areas might flow beyond the focal census tract, or might not typically be captured by a sufficient proportion of the local population to influence our outcome metrics. Across the BLA studies have reported increased human welfare in frontier areas following deforestation (e.g. Weinhold et al. 2015), while other studies from the region (e.g. Pokorny et al. 2013; Ferreira et al. 2010) have contested the view that large scale agricultural developments have benefited human welfare. For example, Garrett et al. (2017) found that low income land uses such as livestock production remained popular in the Eastern Amazon because of limited supply chain mechanisms, but also social preferences and because more importance was given to non-monetary attributes of rural lifestyle (e.g. safety, tranquillity and community relations). Our study indicates that the proposed development of the BLA by agricultural businesses might not uniformly lead to better living conditions locally. At least in some areas, alternatives such as SU PAs that preserve forest cover might perform equally well in offering livelihoods.

Met opmerkingen [b12]: Example to clarify point made

We show that SU PAs reduce deforestation in comparison to mining concessions, but at the expense of locally reduced incomes and higher inequality. Whilst SU PAs did not adversely impact other poverty metrics, there are clear trade-offs between environmental protection and other development goals if establishing SU PAs prevents access to economic benefits of mining activities. We also find that mining impacts stretch beyond the mining site (up to 50 km for deforestation and up to 10 km for income). However, our study only measures relatively short-term impacts and does not take into account health, which is often negatively affected by mining (Von der Golz et al. 2018). Importantly, we are only able to assess legal mining concession even though illegal mining is a well-known issue in the Amazon (Asner & Tupayachi 2016). Further studies should focus on effects of mining different materials and assess effects of illegal mining. Due to demand for minerals mining activities for a variety of mined resources are expected to increase (Ali et al. 2017). Because PAs often overlap with areas of mining interest (Durán et al. 2013; El Bizri et al. 2016), the risk of PADDD events warrants research on how deforestation impacts by mining could be curbed (Sonter et al. 2018). (Asner & Tupayachi 2016).

Our study paves the way for improving future assessments of PA effects. First, comparisons between land uses could be expanded with comparisons across alternative forest-based interventions. For instance, a study in Mexico compared effects of PES programs and PAs, and found that both interventions reduced equal amounts of deforestation, but that PES performed slightly better at poverty alleviation (Sims & Alix-Garcia 2017). In the BLA multiple interventions were implemented that reduced deforestation rates (Arima et al. 2014; Soares-Filho et al. 2014). These include blacklisting of districts with high deforestation rates (Cisneros et al. 2015), declaration of rural environmental registries (Portuguese acronym CAR) to improve monitoring in compliance with the Brazilian Forest Code (L'Roe et al. 2016), and the soy and beef moratoria (an agreement among major beef and soy traders to not buy soy or beef

Met opmerkingen [b13]: Example to clarify point made from land that was cleared after 2006; Gibbs et al. 2015). Comparing interventions could give insight which interventions worked best in achieving environmental and social outcomes.

Secondly, our analysis assumes homogeneity within each treatment and control group whereas management of PAs and land use decisions by individual landholders could have great effects on both forest cover and poverty alleviation (Pringle 2017; Runting et al. 2019). Future studies should pay greater attention to how variations in PA management (but also in control groups) influence outcomes.

Finally, although our study design is specific to the case of the BLA region, agriculture and mining are important competitors for land use worldwide (Golden Kroner et al. 2019). Hence, our findings indicate that PA impact evaluations in other regions could also benefit from elucidating relative PA effects to competing land uses.

3.5 Methods

3.5.1 Unit of analysis

Our analysis focuses on 5,545 CTs within the Brazilian Legal Amazon (BLA) derived from census tracts identified in Brazil's 2,000 human population census. We merged CTs that were smaller than 50 km^2 (mean size 3.5 km2 ± SE; total area covered is less than 1% of the BLA area.) as the boundaries of these census tracts are tightly delimiting villages from the surrounding landscape with which they regularly interact. This decoupling could reduce the ability to detect impacts of land-use decision on the residents of these small census tracts. Our recursive algorithm selected the smallest CT, then merged it with the neighbouring CT situated in the same municipality with whom it shares the longest border – although in practice most of the small census tracts are surrounded by much larger census tracts. The algorithm repeated this procedure until the area of the smallest remaining CT was larger than 50 km^2.

Boundaries of some census tracts changed between 2000 and 2010. We thus followed the approach of Cavatassi et al. (2004) and Andam et al. (2010) and consistently reconstructed data to the 2000 census tract boundaries by first overlapping the 2010 census tract boundaries with the 2000 census tract boundaries and then calculating a weighted average of the 2010 metric for the 2000 census tract boundary. Agricultural census data also experienced boundary changes between 2000 and 2006 and the same procedure was followed to calculate a 2006 metric allocated to the 2000 census tract boundaries.

3.5.2 Metrics

(i) Forest cover loss metrics

We estimated forest cover loss estimates based on the Brazilian PRODES (INPE 2017) and the global forest loss datasets from the University of Maryland (Hansen et al. 2013). The PRODES dataset quantifies large-scale deforestation of primary forest, but ignores loss of secondary forest. Whilst the Hansen dataset combines loss of primary and secondary forest it also includes loss vegetation cover that does not comprise natural forest, such as plantations (Hansen et al. 2013; Tropek et al. 2014). Despite these differences in methodology, both datasets agree on the general spatial patterns and temporal trends in deforestation in BLA (Tyukavina et al. 2017). We chose to use the PRODES dataset in our main analysis because it is widely used in assessment of deforestation in the Amazon (Soares-Filho et al. 2010), and forms the basis of the newly developed MapBiomas dataset (MapBiomas 2015). A robustness check using the Hansen dataset to measure forest cover loss largely revealed similar patterns (Fig. S3.1C) but a small number of minor discrepancies are mentioned in the main text.

(ii) Poverty metrics

We measure four indicators of poverty. We chose these because they are commonly used poverty metrics for assessing impacts of protected areas (e.g. Andam et al. 2010; Canavire-Bacarreza & Hanauer 2013), and comprise both income-based metrics and key dimensions that contribute to multi-dimensional poverty indicators. We do not, however, generate multi-dimensional indicators because we only estimate few indicators and poverty indices are critiqued because they compound accuracy errors (Ravallion, 2010). Two indicators are based on income: i) average income per household head in Brazilian Reais (R$) (corrected for inflation using the INPC/IPC consumer price index (IBGE 2018)), and ii) GINI coefficient, a well-established indicator to assess inequality (Halpern et al. 2013). In addition, we assessed two indicators to account for other dimensions of poverty: i) percentage of literate household heads as a measure of education, and ii) percentage of households with poor sanitation (i.e. without toilets that drain into a sewage system or septic tank) as a measure of health.

3.5.3 Intervention and control definitions

(i) Defining interventions

We define SU PAs in the BLA using the National Cadaster of Protected Areas (Cadastro Nacional de Unidades de Conservação, CNUC) of the Brazilian Ministry of the Environment.

In accordance with Brazilian nomenclature, we classify national forests, extractive reserves, and sustainable development reserves as “sustainable use” (SU) areas (similar to IUCN categories V-VI; Silva, 2005; Figure S3.6A). We thus excluded strictly protected areas and indigenous territories from the analysis, and only focuses on PAs established between 2000 and 2010. We defined protected census tracts using two separate definitions: those with i) at least 10% of their area overlapping with a PA (due to the frequent use of this threshold in other studies e.g. Andam et al. 2010; Hanauer and Canavire-Bacarreza, 2015) and ii) at least 50% of the census tract being protected (a higher threshold that is also used in the same studies). Census tracts with <1% of their area protected were defined as non-protected to ensure a clear distinction in the amount of protection between control and treatment CTs. To further ensure a clear contrast between treatment and control units we excluded protected CTs that contained more than 10% settled land and protected CTs that included a mine.

(ii) Choice of control land uses

We compared PA impacts against two major competitors for land use: agriculture and mining. We chose these two land uses, because they are globally key competitors with PAs for land (Jones et al. 2018; Sonter et al. 2018; Golden Kroner et al. 2019). Moreover, they are the two major drivers of deforestation in the BLA (Nepstad et al. 2014b; Sonter et al. 2017). We did not define forest concessions as an additional type of alternative land use due to their rarity in the BLA (Keller et al. 2007), but recognise remote areas, i.e. census tracts that are not protected and less than 10% of their area is occupied by agricultural areas, as an additional type of land-use.

(iii) Defining controls - agriculture

Firstly, we compare PA impacts against the impacts of agricultural actors. We defined a CT as an agricultural control if at least 10% of the CT is settled by agricultural properties. Following Godar et al. (2016), we defined agricultural controls based on property size of landholders. Two major approaches to define agricultural controls by the size of landholdings can be distinguished: i) the most frequent landholding size category and ii) the landholding size category which occupies the largest proportion of land.. We follow the first approach to ensure that landholding size definitions reflect the type of land-holding of residents that contribute disproportionately to our poverty metrics.

We reclassified the original data into four separate categories: 1) properties smaller than 10 ha are considered "very small" 2) properties between 10 and 50 ha are "small", 3) properties between 50 and 200 ha are "medium", and 4) properties bigger than 200 ha are "large". We determined actor dominance by the category that contains the most properties per census tract (Figure S3.6B). Our classification largely follows global assessments of farm size (Lowder et al. 2016; Samberg et al. 2016), but is adjusted for the Brazilian Amazon because farm sizes in Latin-America are generally bigger than the global average. Our classification is able to tease apart differences in agricultural property size at lower scales than previous assessments conducted in the Amazon that generally treat all properties smaller than 100 ha as smallholder agriculture (e.g. L'Roe et al. 2016). Teasing apart differences in property size below this threshold could be important because 100 ha is substantially higher than the size of the average family farm in Brazil (18.37 ha; Medina et al. 2015).

Furthermore, we distinguished between cropland and pasture (i.e. livestock). Firstly, we calculated the ratio of cropland and pastureland in the agricultural controls of our post-matched sample (Figure S3.4B). Secondly, we performed an analysis that only included agricultural control CTs that are dominated by pasture (>70% of the agricultural area dominated by pasture), but this additional analysis generated similar findings to our main analysis (Figures S3.1B & S3.3). Because of small sample sizes (<5% of the agricultural CTs in the BLA consists of cropland areas dominated by medium or large landholders) we were unable to perform a separate analysis using only cropland agricultural controls.

(iv) Defining controls - mining

Large tracts of the Brazilian Amazon are registered as under consideration for mining, including parts of PAs. However, far fewer areas have been approved and licensed for mining activities. We followed previous studies (Ferreira et al. 2014; Sonter et al. 2017) and only included those mines classified in the Sistema de Informações Geográficas da Mineração (SIGMINE) that have been officially approved and licensed for mining by the Brazilian Minister of Mines and Energy (DNPM 2012). These categories included "*concessão de lavra*", "*concessão de lavra garimpeira*", "*licenciamento*", and "*registro de extração*". We identified mining controls based on the presence/absence of at least one licensed minewithin the census tract (Fig. S3.6C). We only included mines that were licensed during the study period to guarantee that controls at baseline were similar to protected units. We excluded CTs with mines that were licencsed before 2000. Although mining activities might not take place directly after licensing, in many cases infrastructure development occurs before the final approval of the mine and this infrastructure influences deforestation rates (Sonter et al. 2017). To test the robustness of the findings we drew buffers around each mining site and treated all CTs within 5 km, 10 km, 25 km, 50 km and 75 km of the mining site as treated – this matches the maximum reported distance from a mine at which Sonter et al (2007) detected minimizing impacts on deforestation. Due to small sample sizes we were not able to differentiate between various mined resources (for example, gold mining or other polymetallic mines).

3.5.4 Covariates

We selected biophysical and socioeconomic covariates based on their potential to influence the outcome or the relationship between treatment and outcomes (S1; Ferraro et al. 2018b)). These covariates were baseline levels of our poverty measures, slope, size of the census tract, baseline forest cover, agricultural suitability, travel time from the census tract to major cities (pop. size >50,000), population density, and state. For our analyses comparing PAs and agriculture we included presence of mining as a covariate, and we included proportion of settled land as a covariate in our analysis of PAs vs mining.

3.5.5 Matching and post-matching analyses

We used a matching and regression approach to assess the differences in poverty and deforestation outcomes between SU PAs, agriculture and mining. Matching is a data pre-processing technique to optimize the balance of covariates across treated and control units, and is useful when imbalance between treatment and control hinders causal inference (Ho et al. 2007). We used genetic matching with replacement, which uses a machine learning algorithm to select matches and performs well when covariates have skewed distributions (Diamond & Sekhon 2013).

We performed all of our statistical analyses in *R* version 3.3.3 (R Core Team 2018) using the “Matchit” package (Ho et al. 2007). We used a calliper of 1.5 to produce sufficient matches while maintaining matching balance (Figures S3.7-S.313 & Tables S3.4-S3.10). We used post-matching standardized mean differences of <0.25 as an acceptable balance between treatment and control groups for each covariate (Stuart 2010). Finally, we conducted an Ordinary Least Squares (OLS) regression on the matched sample using weights obtained from the matching to adjust for any imbalances in covariate distributions.

We first conducted an impact evaluation and compared the average impact of protection on deforestation and poverty against a counterfactual without distinguishing the type of land use in the controls. We then split our control group and compared the relative effects of SU PAs on forest loss and poverty indicators against our different control groups (agriculture according to property size, and mining). For the comparisons with the agricultural groups we compared the proportion of settled land post-matching to confirm that the amount of settled land did not confound our main findings (Fig. S3.4A). We conducted robustness tests to check for spatial autocorrelation. Although spatial autocorrelation is present within our models, our estimated value of spatial autocorrelation (Moran's I) is very low so we do not expect spatial autocorrelation to interfere with our main findings (Table S3.3).

3.6 References

Ali, S.H., Giurco, D., Arndt, N., Nickless, E., Brown, G., Demetriades, A., Durrheim, R., Enriquez, M.A., Kinnaird, J., Littleboy, A., Meinert, L.D., Oberhänsli, R., Salem, J., Schodde, R., Schneider, G., Vidal, O. & Yakovleva, N. (2017). Mineral supply for sustainable development requires resource governance. *Nature*, 543, 367–372.

Andam, K.S., Ferraro, P.J., Sims, K.R.E., Healy, A. & Holland, M.B. (2010). Protected areas reduced poverty in Costa Rica and Thailand. *Proc. Natl. Acad. Sci. U. S. A.*, 107, 9996–10001.

Arima, E.Y., Barreto, P., Araújo, E. & Soares-Filho, B. (2014). Public policies can reduce tropical deforestation: Lessons and challenges from Brazil. *Land use policy*, 41, 465–473.

Asner, G.P. & Tupayachi, R. (2016). Accelerated losses of protected forests from gold mining in the Peruvian Amazon. *Environ. Res. Lett.*, 12, 094004.

Austin, K.G., Schwantes, A., Gu, Y. & Kasibhatla, P.S. (2019). What causes deforestation in Indonesia? *Environ. Res. Lett.*, 14, 024007.

Barber, C.P., Cochrane, M.A., Souza, C.M. & Laurance, W.F. (2014). Roads, deforestation, and the mitigating effect of protected areas in the Amazon. *Biol. Conserv.*, 177, 203–209.

Barrett, C.., Reardon, T. & Webb, P. (2001). Nonfarm income diversification and household livelihood strategies in rural Africa: concepts, dynamics, and policy implications. *Food Policy*, 26, 315–331.

Bebbington, A., Hinojosa, L., Bebbington, D.H., Burneo, M.L. & Warnaars, X. (2008). Contention and Ambiguity: Mining and the Possibilities of Development. *Dev. Change*, 39, 887–914.

El Bizri, H.R., Macedo, J.C.B., Paglia, A.P. & Morcatty, T.Q. (2016). Mining undermining Brazil's environment. *Science*, 353, 228.

Canavire-Bacarreza, G. & Hanauer, M.M. (2013). Estimating the Impacts of Bolivia's Protected Areas on Poverty. *World Dev.*, 41, 265–285.

Cisneros, E., Zhou, S.L. & Börner, J. (2015). Naming and Shaming for Conservation:

Evidence from the Brazilian Amazon. *PLoS One*, 10, e0136402.

Diamond, A. & Sekhon, J.S. (2013). Genetic matching for estimating causal effects: A general multivariate matching method for achieving balance in observational studies. *Rev. Econ. Stat.*, 95, 932–945.

DNPM, D.N. de P.M. (2012). Processos Minerarios: Sistema de Informações Geográficas da Mineração (SIGMINE) [WWW Document]. URL http://sigmine.dnpm.gov.br/

Durán, A.P., Rauch, J. & Gaston, K.J. (2013). Global spatial coincidence between protected areas and metal mining activities. *Biol. Conserv.*, 160, 272–278.

Edwards, D.P., Sloan, S., Weng, L., Dirks, P., Sayer, J. & Laurance, W.F. (2014). Mining and the African Environment. *Conserv. Lett.*, 7, 302–311.

Escobar, H. (2019). Bolsonaro's first moves have Brazilian scientists worried. *Science*, 363, 330.

Ferraro, P.J., Sanchirico, J.N. & Smith, M.D. (2018a). Causal inference in coupled human and natural systems. *Proc. Natl. Acad. Sci. U. S. A.*, 201805563.

Ferraro, P.J., Sanchirico, J.N. & Smith, M.D. (2018b). Causal inference in coupled human and natural systems. *Proc. Natl. Acad. Sci. U. S. A.*, 201805563.

Ferreira, F.H.G., Leite, P.G. & Ravallion, M. (2010). Poverty reduction without economic growth?: Explaining Brazil's poverty dynamics, 1985–2004. *J. Dev. Econ.*, 93, 20–36.

Ferreira, J., Aragão, L.E.O.C., Barlow, J., Barreto, P., Berenguer, E., Bustamante, M., Gardner, T.A., Lees, A.C., Lima, A., Louzada, J., Pardini, R., Parry, L., Peres, C.A., Pompeu, P.S., Tabarelli, M. & Zuanon, J. (2014). Environment and Development. Brazil's environmental leadership at risk. *Science*, 346, 706–7.

Foley, J.A., Defries, R., Asner, G.P., Barford, C., Bonan, G., Carpenter, S.R., Chapin, F.S., Coe, M.T., Daily, G.C., Gibbs, H.K., Helkowski, J.H., Holloway, T., Howard, E.A., Kucharik, C.J., Monfreda, C., Patz, J.A., Prentice, I.C., Ramankutty, N. & Snyder, P.K. (2005). Global consequences of land use. *Science*, 309, 570–4.

Garrett, R.D., Gardner, T.A., Morello, T.F., Marchand, S., Barlow, J., Ezzine de Blas, D., Ferreira, J., Lees, A.C. & Parry, L. (2017). Explaining the persistence of low income and environmentally degrading land uses in the Brazilian Amazon. Ecol. Soc., 22, art27.

Gibbs, H.K., Rausch, L., Munger, J., Schelly, I., Morton, D.C., Noojipady, P., Soares-Filho, B., Barreto, P., Micol, L. & Walker, N.F. (2015). Environment and development. Brazil's Soy Moratorium. *Science*, 347, 377–8.

Godar, J., Gardner, T.A., Tizado, E.J. & Pacheco, P. (2014). Actor-specific contributions to the deforestation slowdown in the Brazilian Amazon. *Proc. Natl. Acad. Sci. U. S. A.*, 111, 15591–6.

Golden Kroner, R.E., Qin, S., Cook, C.N., Krithivasan, R., Pack, S.M., Bonilla, O.D., Cort-Kansinally, K.A., Coutinho, B., Feng, M., Martínez Garcia, M.I., He, Y., Kennedy, C.J., Lebreton, C., Ledezma, J.C., Lovejoy, T.E., Luther, D.A., Parmanand, Y., Ruíz-Agudelo, C.A., Yerena, E., Morón Zambrano, V. & Mascia, M.B. (2019). The uncertain future of protected lands and waters. *Science*, 364, 881–886.

von der Goltz, J. & Barnwal, P. (2018). Mines: The local wealth and health effects of mineral mining in developing countries. *J. Dev. Econ.*

Halpern, B.S., Klein, C.J., Brown, C.J., Beger, M., Grantham, H.S., Mangubhai, S., Ruckelshaus, M., Tulloch, V.J., Watts, M., White, C. & Possingham, H.P. (2013). Achieving the triple bottom line in the face of inherent trade-offs among social equity, economic return, and conservation. *Proc. Natl. Acad. Sci. U. S. A.*, 110, 6229–6234.

Hansen, M.C., Potapov, P. V, Moore, R., Hancher, M., Turubanova, S.A., Tyukavina, A., Thau, D., Stehman, S. V, Goetz, S.J., Loveland, T.R., Kommareddy, A., Egorov, A., Chini, L., Justice, C.O. & Townshend, J.R.G. (2013). High-Resolution Global Maps of 21st-Century Forest Cover Change. *Science (80-.).*, 342, 850–853.

Hecht, S.B. (1985). Environment, development and politics: Capital accumulation and the livestock sector in Eastern Amazonia. *World Dev.*, 13, 663–684.

Heilmayr, R. & Lambin, E.F. (2016). Impacts of nonstate, market-driven governance on Chilean forests. *Proc. Natl. Acad. Sci. U. S. A.*, 113, 2910–5.

Ho, D.E., Imai, K., King, G. & Stuart, E.A. (2007). Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. *Polit. Anal.*, 15, 199–236.

IBGE, I.B. de G. e E. (2018). Índice Nacional de Preços ao Consumidor Amplo - IPCA e Índice Nacional de Preços ao Consumidor - INPC [WWW Document]. URL https://ww2.ibge.gov.br/home/estatistica/indicadores/precos/inpc_ipca

INPE, I.N. de P.E. (2017). PRODES — Monitoramento da Floresta Amazônica Brasileira por Satélite [WWW Document]. URL http://www.obt.inpe.br/prodes

IUCN & UNEP-WCMC. (2016). The World Database on Protected Areas (WDPA).

Jones, K.R., Venter, O., Fuller, R.A., Allan, J.R., Maxwell, S.L., Negret, P.J. & Watson, J.E.M. (2018). One-third of global protected land is under intense human pressure. *Science*, 360, 788–791.

Joppa, L.N. & Pfaff, A. (2009). High and Far: Biases in the Location of Protected Areas. *PLoS One*, 4.

Keller, M., Asner, G.P., Blate, G., McGlocklin, J., Merry, F., Peña-Claros, M. & Zweede, J. (2007). Timber production in selectively logged tropical forests in South America. *Front. Ecol. Environ.*, 5, 213–216.

Kere, E.N., Choumert, J., Combes Motel, P., Combes, J.L., Santoni, O. & Schwartz, S. (2017). Addressing Contextual and Location Biases in the Assessment of Protected Areas Effectiveness on Deforestation in the Brazilian Amazônia. *Ecol. Econ.*, 136, 148–158.

L'Roe, J., Rausch, L., Munger, J. & Gibbs, H.K. (2016). Mapping properties to monitor forests: Landholder response to a large environmental registration program in the Brazilian Amazon. *Land use policy*, 57, 193–203.

Loayza, N. & Rigolini, J. (2016). The Local Impact of Mining on Poverty and Inequality: Evidence from the Commodity Boom in Peru. World Dev., 84, 219–234.

Lowder, S.K., Skoet, J. & Raney, T. (2016). The Number, Size, and Distribution of Farms, Smallholder Farms, and Family Farms Worldwide. *World Dev.*, 87, 16–29.

Malm, O. (1998). Gold Mining as a Source of Mercury Exposure in the Brazilian Amazon. *Environ. Res.*, 77, 73–78.

MapBiomas. (2015). Project MapBiomas — Collection 2 of Brazilian Land Cover & Use Map Series [WWW Document]. URL http://mapbiomas.org/

Medina, G., Almeida, C., Novaes, E., Godar, J. & Pokorny, B. (2015). Development Conditions for Family Farming: Lessons From Brazil. *World Dev.*, 74, 386–396.

Meyfroidt, P. (2018). Trade-offs between environment and livelihoods: Bridging the global land use and food security discussions. *Glob. Food Sec.*, 16, 9–16.

Naeem, S., Chazdon, R., Duffy, J.E., Prager, C. & Worm, B. (2016). Biodiversity and human well-being: an essential link for sustainable development. *Proc. R. Soc. B Biol. Sci.*, 283, 20162091.

Nepstad, D., McGrath, D., Stickler, C., Alencar, A., Azevedo, A., Swette, B., Bezerra, T., DiGiano, M., Shimada, J., Seroa da Motta, R., Armijo, E., Castello, L., Brando, P., Hansen, M.C., McGrath-Horn, M., Carvalho, O. & Hess, L. (2014a). Slowing Amazon deforestation through public policy and interventions in beef and soy supply chains. *Science*, 344, 1118–23.

Nepstad, D., McGrath, D., Stickler, C., Alencar, A., Azevedo, A., Swette, B., Bezerra, T., DiGiano, M., Shimada, J., Seroa da Motta, R., Armijo, E., Castello, L., Brando, P., Hansen, M.C., McGrath-Horn, M., Carvalho, O. & Hess, L. (2014b). Slowing Amazon deforestation through public policy and interventions in beef and soy supply chains. *Science*, 344, 1118–23.

Nepstad, D., Soares-Filho, B.S., Merry, F., Lima, A., Moutinho, P., Carter, J., Bowman, M., Cattaneo, A., Rodrigues, H., Schwartzman, S., McGrath, D.G., Stickler, C.M., Lubowski, R., Piris-Cabezas, P., Rivero, S., Alencar, A., Almeida, O. & Stella, O. (2009). Environment. The end of deforestation in the Brazilian Amazon. *Science*, 326, 1350–1.

Nolte, C., Agrawal, A., Silvius, K.M. & Soares-Filho, B.S. (2013). Governance regime and location influence avoided deforestation success of protected areas in the Brazilian Amazon. *Proc. Natl. Acad. Sci. U. S. A.*, 110, 4956–4961.

Pacheco, P. (2009). Agrarian Reform in the Brazilian Amazon: Its Implications for Land Distribution and Deforestation. *World Dev.*, 37, 1337–1347.

Pack, S.M., Ferreira, M.N., Krithivasan, R., Murrow, J., Bernard, E. & Mascia, M.B. (2016). Protected area downgrading, downsizing, and degazettement (PADDD) in the Amazon. *Biol. Conserv.*, 197, 32–39.

Pokorny, B., de Jong, W., Godar, J., Pacheco, P. & Johnson, J. (2013). From large to small: Reorienting rural development policies in response to climate change, food security and poverty. *For. Policy Econ.*, 36, 52–59.

Pringle, R.M. (2017). Upgrading protected areas to conserve wild biodiversity. *Nature*, 546, 91–99.

Richards, P.D. & VanWey, L. (2016). Farm-scale distribution of deforestation and remaining forest cover in Mato Grosso. *Nat. Clim. Chang.*, 6, 418–425.

Richards, P.D., Walker, R.T. & Arima, E.Y. (2014). Spatially complex land change: The Indirect effect of Brazil's agricultural sector on land use in Amazonia. *Glob. Environ. Chang.*, 29, 1–9.

Rodrigues, A.S.L., Andelman, S.J., Bakarr, M.I., Boitani, L., Brooks, T.M., Cowling, R.M., Fishpool, L.D.C., da Fonseca, G.A.B., Gaston, K.J., Hoffmann, M., Long, J.S., Marquet, P.A., Pilgrim, J.D., Pressey, R.L., Schipper, J., Sechrest, W., Stuart, S.N., Underhill, L.G., Waller, R.W., Watts, M.E.J. & Yan, X. (2004). Effectiveness of the global protected area network in representing species diversity. *Nature*, 428, 640–643.

Runting, R.K., Ruslandi, Griscom, B.W., Struebig, M.J., Satar, M., Meijaard, E., Burivalova, Z., Cheyne, S.M., Deere, N.J., Game, E.T., Putz, F.E., Wells, J.A., Wilting, A., Ancrenaz, M., Ellis, P., Khan, F.A.A., Leavitt, S.M., Marshall, A.J., Possingham, H.P., Watson, J.E.M. & Venter, O. (2019). Larger gains from improved management over sparing–sharing for tropical forests. *Nat. Sustain.*, 2, 53–61.

Samberg, L.H., Gerber, J.S., Ramankutty, N., Herrero, M. & West, P.C. (2016). Subnational distribution of average farm size and smallholder contributions to global food production. *Environ. Res. Lett.*, 11, 124010.

Schneider, M. & Peres, C.A. (2015). Environmental Costs of Government-Sponsored Agrarian Settlements in Brazilian Amazonia. *PLoS One*, 10, e0134016.

Soares-Filho, B., Moutinho, P., Nepstad, D., Anderson, A., Rodrigues, H., Garcia, R., Dietzsch, L., Merry, F., Bowman, M., Hissa, L., Silvestrini, R. & Maretti, C. (2010). Role of Brazilian Amazon protected areas in climate change mitigation. *Proc. Natl. Acad. Sci. U. S. A.*, 107, 10821–6.

Soares-Filho, B., Rajão, R., Macedo, M., Carneiro, A., Costa, W., Coe, M., Rodrigues, H. & Alencar, A. (2014). Land use. Cracking Brazil's Forest Code. *Science*, 344, 363–4.

Socolar, J.B., Valderrama Sandoval, E.H. & Wilcove, D.S. (2019). Overlooked biodiversity loss in tropical smallholder agriculture. *Conserv. Biol.*, cobi.13344.

Sonter, L.J., Ali, S.H. & Watson, J.E.M. (2018). Mining and biodiversity: key issues and research needs in conservation science. *Proc. R. Soc. B Biol. Sci.*, 285, 20181926.

Sonter, L.J., Herrera, D., Barrett, D.J., Galford, G.L., Moran, C.J. & Soares-Filho, B.S. (2017). Mining drives extensive deforestation in the Brazilian Amazon. *Nat. Commun.*, 8, 1013.

Stuart, E.A. (2010). Matching Methods for Causal Inference: A Review and a Look Forward. *Stat. Sci.*, 25, 1–21.

Team, R.C. (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2012. *URL http//www.R-project.org*.

Tropek, R., Sedláček, O., Beck, J., Keil, P., Musilová, Z., Símová, I. & Storch, D. (2014). Comment on "High-resolution global maps of 21st-century forest cover change". *Science*, 344, 981.

Tyukavina, A., Hansen, M.C., Potapov, P. V., Stehman, S. V., Smith-Rodriguez, K., Okpa, C. & Aguilar, R. (2017). Types and rates of forest disturbance in Brazilian Legal Amazon, 2000–2013. *Sci. Adv.*, 3, e1601047.

Weinhold, D., Reis, E.J. & Vale, P.M. (2015). Boom-bust patterns in the Brazilian Amazon. *Glob. Environ. Chang.*, 35, 391–399.

3.7 Supplementary information

S1 Confounding factors

Our selection of covariates follows other research on protected areas (Andam et al. 2010; Canavire-Bacarreza & Hanauer 2013). Table S3.1 presents the confounding factors used in our analysis, and the rationale for including them. We also included forest and poverty baselines in our models as covariates to match on.

Table S3.2 presents baseline levels of covariates for our treatment and control groups. As expected protected units show higher baseline tree cover than the unprotected controls. Protected units also show lower baseline income and lower baseline literacy levels, but inequality is also lower. In addition, PAs are established in areas with lower population density, lower agricultural suitability, longer travel time to major cities, and larger census tracts.

When we tease apart our control group we observe significant heterogeneity within the control group. As expected areas dominated by smallholders show higher tree cover, lower average income, and lower literacy levels at baseline than areas dominated by larger-sized landholders. However, inequality in these areas is also lower than in the other agricultural controls. Values of other covariates differ substantially within each agricultural control category.

Mining sites show the lowest tree cover at baseline in comparison to other controls. Baseline levels of income, literacy, and inequality are higher in mining sites than in the homogeneous controls. Mining sites are also associated with higher population densities, less steep slopes, and slightly larger census tracts, but covariates differ.

Table S3.1. Confounding factors used in analyses. The unit of analysis is the census tract (CT).

Confounding factor	Definition and data source	Rationale for inclusion
Area	Size of the census tract (km^2). Data: CNUC	Size of the administrative area s linked to poverty outcomes (Andam et al. 2010)
Slope	Slope (degree), extracted from data provided by the International Institute for Applied Systems Analysis (Fischer et al. 2007)	Slope influences the distributions of PAs and people's livelihood decisions (Joppa & Pfaff 2009; Gentle & Maraseni 2012)
Population density	Population density year 2000 (IBGE)	Population pressure affects poverty outcomes, resource use and PA allocation (Wittemyer et al. 2008)
Agricultural suitability	Weighted agricultural suitability based on soil, slope and floodability restrictions. Calculations are based on Nepstad (Nepstad et al. 2008)	Agriculture is a leading cause of land-cover change and an important source of livelihoods for local communities (Angelsen 2010)
Proportion settled area	Proportion of settled land in the census tract, based on calculations from agricultural census (IBGE)	Land settlement influences land use choices, people's welfare, and resource use (Hecht 1985)
Mine presence established before 2000	Binary exact match based on the presence/absence of legal mining before 2000 (SIGMINE; DNPM 2012)	Mining is an important land use and influences land use choices, livelihoods, and PA allocation decisions (Ali et al. 2017; Sonter et al. 2018)
Travel time to pop. centre (>50,000 inhabitants)	Travel time to major cities using calculations as in Nolte et al. (2013), based on Soares-Filho et al. (2006)	Access to markets and services (e.g., technical assistance, health centres) significantly affects livelihoods, PA establishment decisions, and other related land-use choices (Angelsen 2010)
State	Exact match based on the state the CT is located in	Federal states have significant autonomy in decision-making and planning that can influence deforestation and development

Table S3.2. Summary statistics of covariates for protected and control units. Reported values are coefficient ± standard error

	SU PAs est. after 2000	Non-PA	Remote (<1% of land settled)	Very small (<10 ha)	Medium (10-50 ha)	Large (50-200 ha)	Very large (>200 ha)	Mining
Baseline tree cover (%)	89.72 ± 13.84	57.35 ± 24.00	75.72 ± 22.14	60.65 ± 21.24	50.44 ± 23.16	52.82 ± 22.69	54.14 ± 23.52	49.55 ± 22.25
Area of census tract (km^2)	2644 ± 4213	410 ± 681	751 ± 1251	306 ± 594	223 ± 283	290 ± 355	593 ± 697	493 ± 657
Slope (degree)	27.4 ± 109.5	11.6 ± 82.7	29.5 ± 135.4	26.6 ± 126.0	22.2 ± 111	3.3 ± 43.4	3.9 ± 34.3	2.9 ± 17.7
Agricultural suitability (0-1)	0.80 ± 0.49	0.63 ± 0.51	0.74 ± 0.54	0.64 ± 0.51	0.70 ± 0.50	0.58 ± 0.49	0.62 ± 0.15	0.57 ± 0.49
Travel time to major cities (minutes)	3989 ± 2696	1249 ± 1115	1956 ± 1838	1016 ± 919	917 ± 676	1085 ± 758	1391 ± 1124	1185 ± 1009
Baseline population density (pop. per km^2)	7.24 ± 33.19	15.62 ± 68.70	20.57 ± 118.90	25.97 ± 72.05	22.59 ± 59.87	12.97 ± 48.06	8.34 ± 56.97	32.66 ± 122.18
Baseline income (Reais)	392 ± 386	633 ± 568	433 ± 329	573 ± 755	617 ± 459	617 ± 400	889 ± 829	775 ± 769
Baseline GINI index	0.36 ± 0.14	0.43 ± 0.13	0.38 ± 0.14	0.40 ± 0.13	0.45 ± 0.12	0.44 ± 0.12	0.46 ± 0.14	0.46 ± 0.12
Baseline literacy (%)	0.53 ± 0.20	0.65 ± 0.16	0.58 ± 0.19	0.61 ± 0.17	0.65 ± 0.13	0.66 ± 0.15	0.70 ± 0.15	0.70 ± 0.14
Baseline poor sanitation (%)	0.96 ± 0.11	0.94 ± 0.13	0.93 ± 0.13	0.90 ± 0.18	0.95 ± 0.10	0.95 ± 0.11	0.92 ± 0.15	0.90 ± 0.17
Sample size	213	3251	480	439	203	1547	582	352

Table S3.3. Spatial autocorrelation results for each control group and outcome.

Outcome	Observed Moran's I (P-value)						
		Agricultural landholder category					
	Non-PA	Remote	Very small	Small	Medium	Large	Mining
Deforestation	-0.001 (0.022)	-0.001 (0.006)	-0.009 (0.871)	-0.020 (0.811)	-0.004 (0.173)	-0.006 (0.351)	-0.006 (0.777)
Income	0.000 (0.004)	0.004 (<0.001)	-0.009 (0.782)	-0.017 (0.460)	-0.004 (0.213)	-0.014 (0.345)	-0.007 (0.783)
Inequality	-0.006 (0.060)	-0.005 (0.913)	-0.009 (0.820)	-0.016 (0.987)	-0.011 (0.613)	-0.008 (0.687)	-0.005 (0.394)
Literature	0.029 (<0.001)	0.031 (<0.001)	0.013 (<0.001)	-0.008 (0.118)	0.037 (<0.001)	0.022 (<0.001)	0.006 (<0.001)
Sanitation	0.007 (<0.001)	0.003 (<0.001)	0.003 (<0.001)	-0.024 (0.949)	0.001 (0.003)	0.005 (<0.001)	0.002 (<0.001)

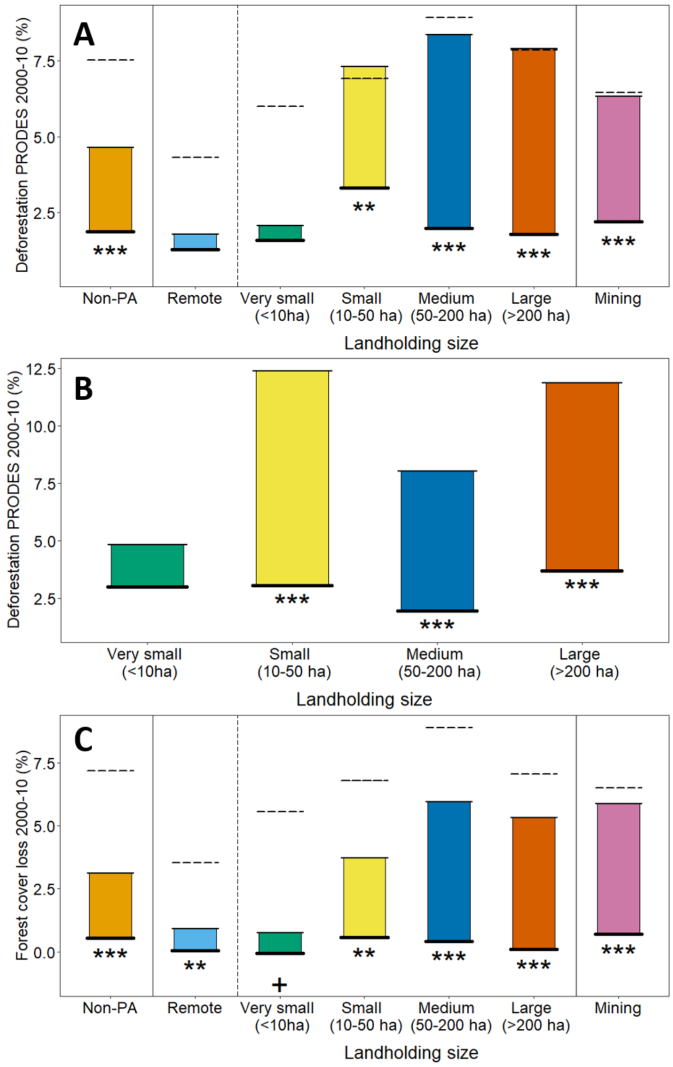

Figure S3.1. Estimated impacts of sustainable use protected area (SU PAs) established after 2000 on forest (A) after raising the threshold from 10% to 50% of census tracts overlapping a SU PA (using the PRODES deforestation dataset), (B) in comparison to pastureland, and (C) using the global forest loss dataset from the University of Maryland, in comparison to non-protection, agricultural controls split by dominant size of landholdings, and registered mining sites, in census tracts (CTs) in the Brazilian Amazon. Non-PA controls include all the CTs with less than 1% of their land protected. Remote areas are defined as areas with less than 10% of the CTs settled. Dominant agricultural landholder size is determined by the category that contains the most properties per CT. Mining sites are defined as CTs with recorded licensed mining activities present after 2000. Thick lines represent treatment i.e. SU PAs, thin lines represent modelled values in controls matched to the treatment across all ten covariates. Dashed lines represent average level in non-matched control group. Significance: *** $P < 0.001$, ** $P < 0.01$, + $P < 0.1$.

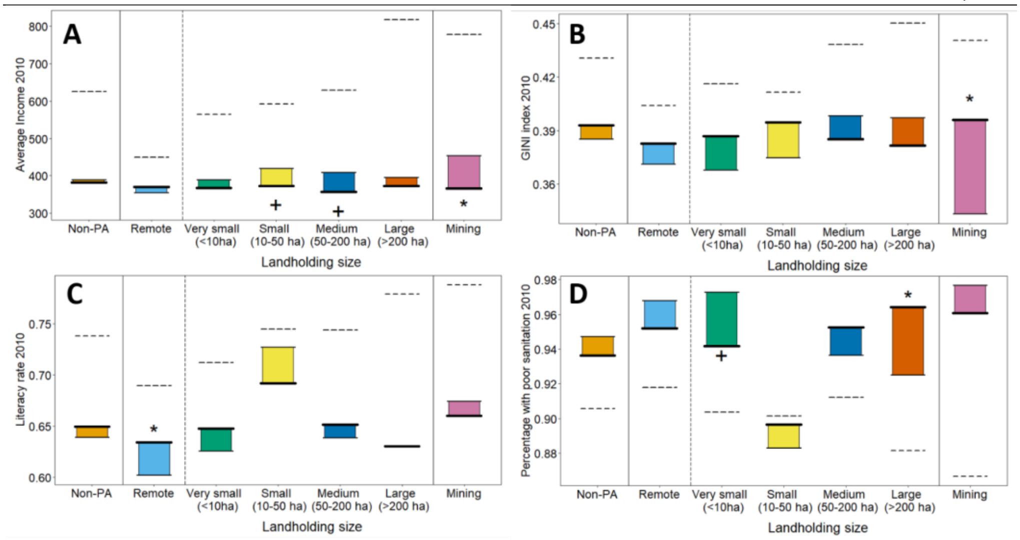

Figure S3.2. Estimated impacts after raising the threshold from 10% to 50% of census tracts overlapping a sustainable use protected area (SU PA) established after 2000 on (A) average income, (B) GINI index, (C) literacy, and (D) sanitation) in comparison to non-protection, agricultural controls split by dominant size of landholdings, and registered mining sites, in census tracts (CTs) of the Brazilian Amazon. Non-PA controls include all the CTs with less than 1% of their land protected. Remote areas are defined as areas with less than 10% of the CT settled. Dominant agricultural landholder size is determined by the category that contains the most properties per CT. Mining sites are defined as CTs with recorded licensed mining activities present after 2000. Thick lines represent treatment i.e. PAs, thin lines represent modelled values in controls matched to the treatment across all ten covariates. Dashed lines represent average level in non-matched control group. Significance: * $P < 0.05$, + $P < 0.1$.

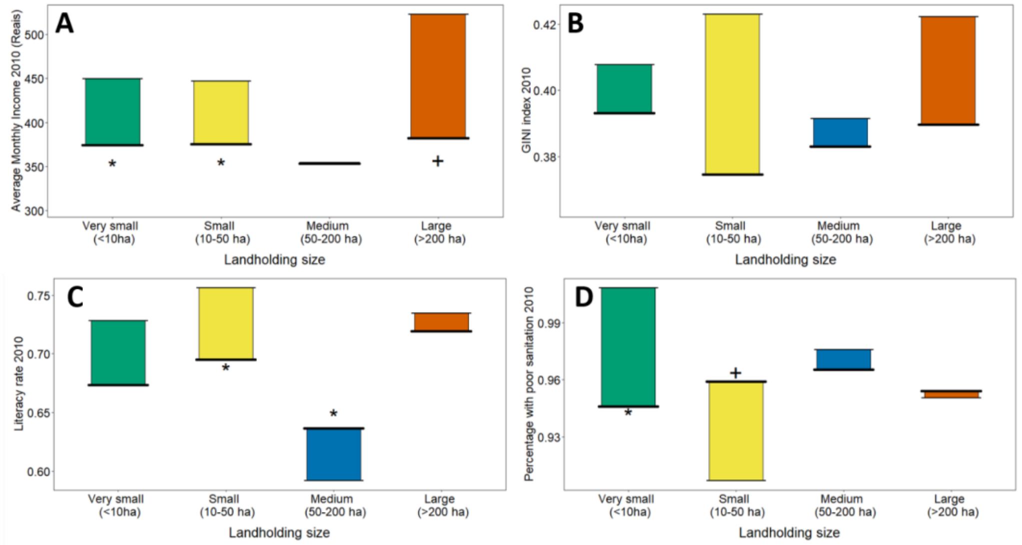

Figure S3.3. Estimated impacts of sustainable use (SU) protected areas established after 2000 on (A) average income, (B) GINI index, (C) literacy, and (D) sanitation in comparison to area dominated by pastureland (>70% of the settled land classified as pastureland) split by dominant size of landholdings, in census tracts (CTs) in the Brazilian Amazon. Dominant pasture landholder size is determined by the category that contains the most properties per CT. Thick lines represent treatment i.e. PAs, thin lines represent modelled values in controls matched to the treatment across all ten covariates. Significance: * $P < 0.05$, + $P < 0.1$.

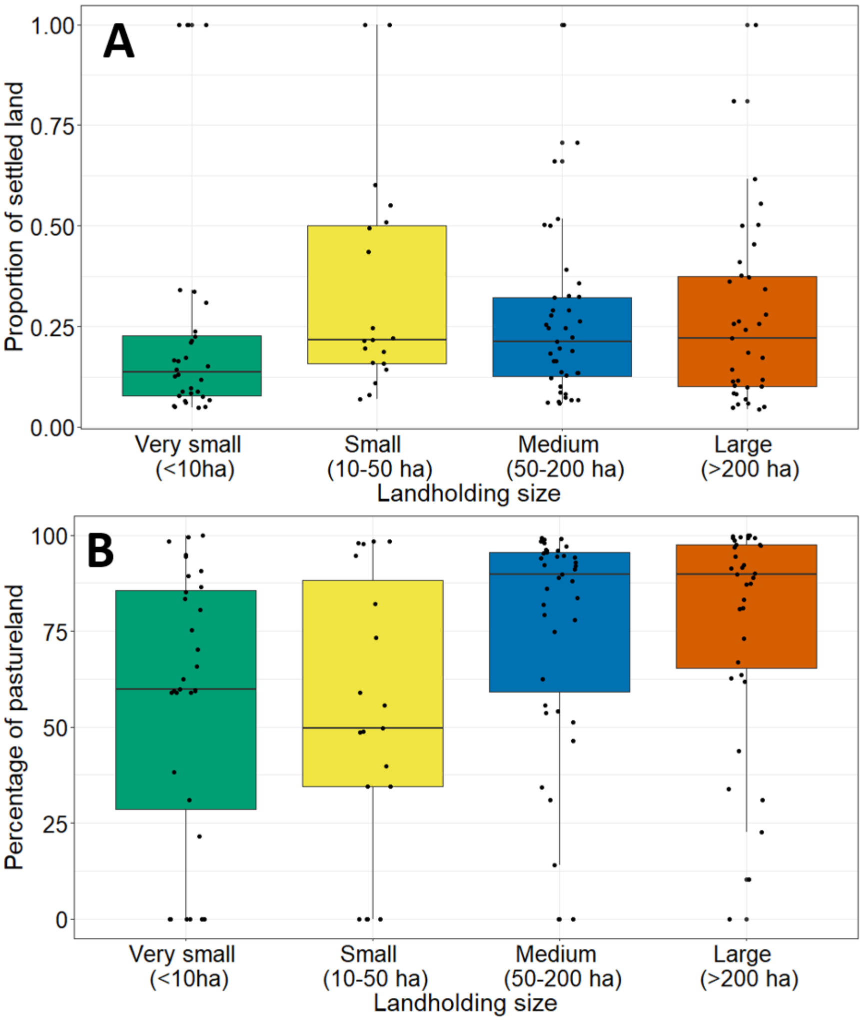

Figure S3.4. Box plots of (A) percentage of settled land in the post-matched sample for each landholder size category, and (B) percentage of pastureland in the post-matched sample for each landholder size category. Percentage of settled land does not differ significantly between landholding size categories ($F = 0.69$, $df = 121$, $P = 0.56$). Percentage of cropland differs significantly between the categories ($F = 0.69$, $df = 121$, $P = 0.56$) - percentage in very small and small categories is smaller than areas dominated by medium ($P = 0.03$) and large ($P = 0.03$) landholders. Dominant landholder size is determined by the category that contains the most properties per CT. Box boundaries represent upper and lower quantiles, horizontal lines inside box represent medians, vertical lines represent range, dots represent individual observations.

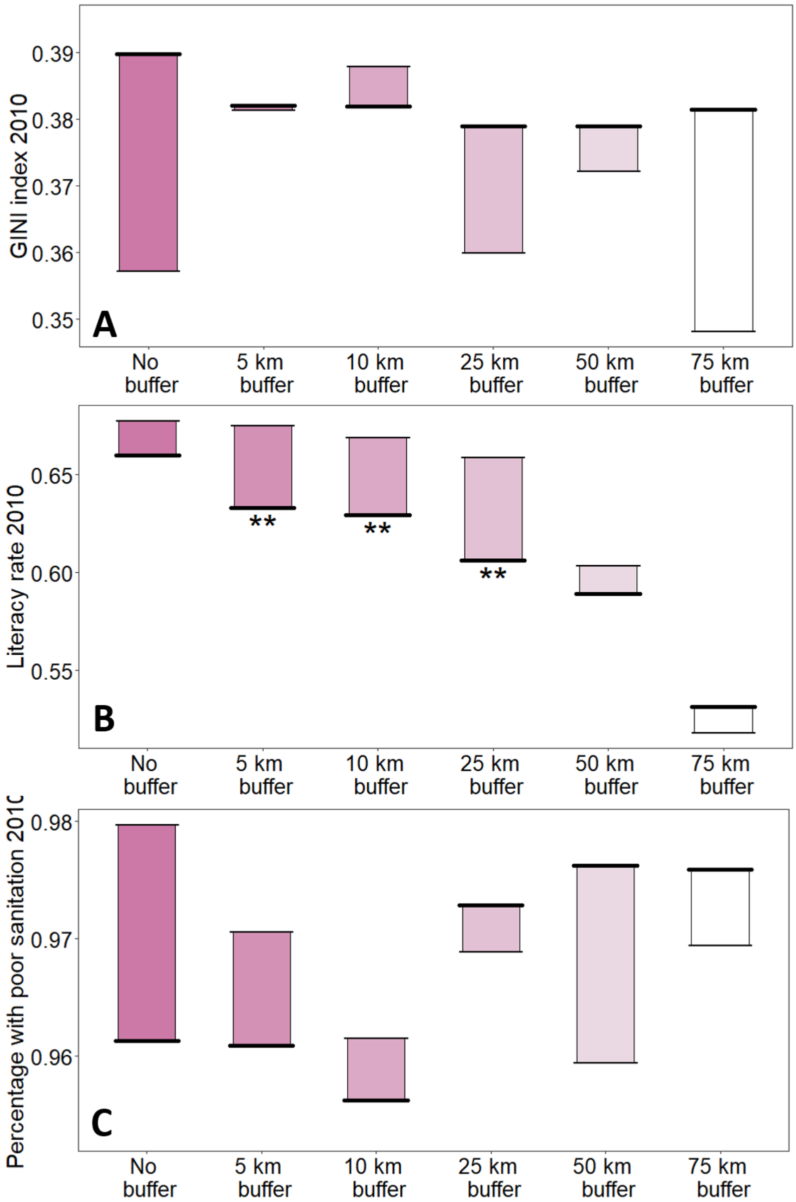

Figure S3.5. Estimated impacts of sustainable use protected areas (SU PAs) established after 2000 on (A) GINI, (B) literacy and (C) sanitation in comparison to mining concessions across different buffer distances, in census tracts of the Brazilian Amazon. Thick lines represent treatment i.e. PAs, thin lines represent modelled values in controls matched to the treatment across all ten covariates. Significance: ** $P < 0.01$.

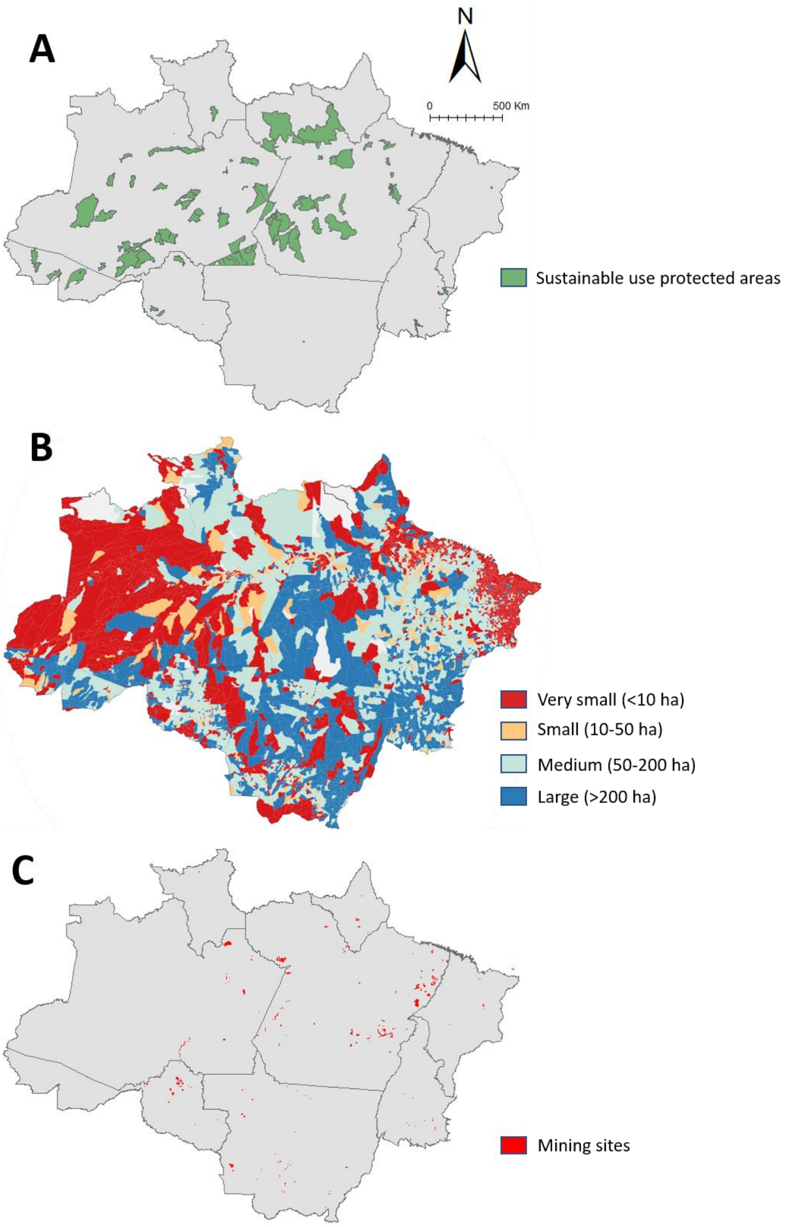

Figure S3.6. Schematic maps of (A) sustainable use protected areas established 2000-2010, (B) dominant landholders, and (C) registered mining sites 2000-2010 in the Brazilian Amazon. Dominant landholder size is determined by the category that contains the most properties per CT.

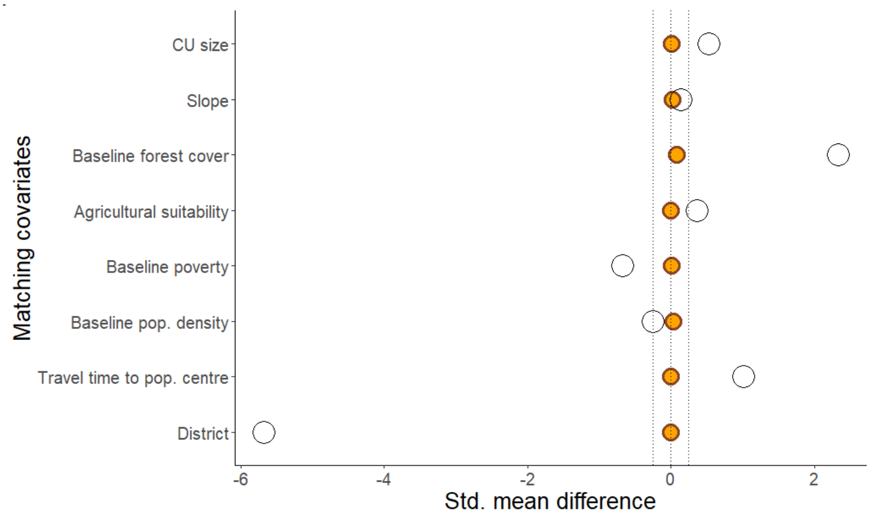

Figure S3.7. Balance of covariates before and after genetic matching, using sustainable use protected areas established after 2000 as treatment, non-protected areas as control and deforestation as outcome variable. Standardized mean difference for confounding covariates before (open circles) and after matching (orange circles). Dotted lines indicate acceptable standard mean difference after matching to reach balance (Stuart 2010).

Table S3.4. Balance of covariates before and after genetic matching, using sustainable use protected areas established after 2000 as treatment, non-protected areas as control and poverty indicators as outcome variables. Analysis was repeated for each poverty indicator separately, but only covariate balance for income outcome is reported. For other poverty indicators only post-matched balance of the poverty indicator is reported.

	Before matching			After matching		
	Means treated (n = 213)	Means control (n = 3250)	Std. mean Diff.	Means treated (n = 128)	Means control (n=186)	Std. mean Diff.
CT size	2644	410	0.530	1822	1847	-0.006
Slope	27.4	11.6	0.144	30.9	30.6	0.002
Baseline forest cover	89.7	57.4	2.339	89.0	87.9	0.074
Agricultural suitability	0.802	0.628	0.359	0.82	0.83	-0.012
Baseline income	393	633	-0.621	364	368	-0.010
Baseline GINI	0.36	0.43	-0.473	0.36	0.36	0.020
Baseline literacy	0.53	0.65	-0.599	0.53	0.53	0.015
Baseline sanitation	0.96	0.94	0.182	0.97	0.97	-0.048
Baseline pop. density	7.2	15.6	-0.253	5.8	5.2	0.019
Travel time to pop. centre	3989	1249	1.016	3601	3539	0.023

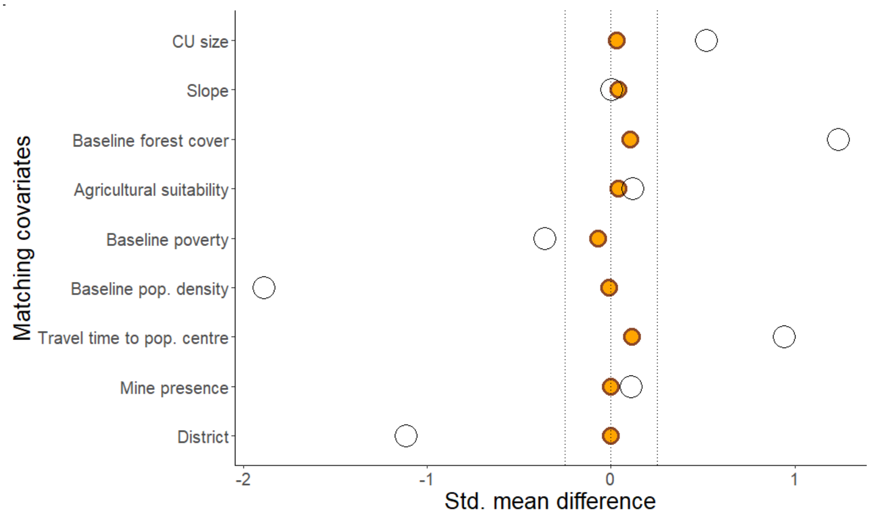

Figure S3.8. Balance of covariates before and after genetic matching, using sustainable use protected areas established after 2000 as treatment, remote non-settled areas as control and deforestation as outcome variable. Standardized mean difference for confounding covariates before (open circles) and after matching (orange circles). Dotted lines indicate acceptable standard mean difference after matching to reach balance (Stuart 2010).

Table S3.5. Balance of covariates before and after genetic matching, using sustainable use protected areas established after 2000 as treatment, non-protected areas as control and poverty indicators as outcome variables. Analysis was repeated for each poverty indicator separately, but only covariate balance for income outcome is reported. For other poverty indicators only post-matched balance of the poverty indicator is reported.

	Before matching			After matching		
	Means treated (n = 176)	Means control (n = 480)	Std. mean Diff.	Means treated (n = 145)	Means control (n = 86)	Std. mean Diff.
CT size	3084	751	0.518	2063	1977	0.019
Slope	29.9	29.5	0.004	30.2	24.9	0.045
Baseline forest cover	91.5	75.7	1.238	90.7	89.7	0.077
Agricultural suitability	0.80	0.74	0.117	0.84	0.80	0.081
Baseline income	386	433	-0.113	332	342	-0.023
Baseline GINI	0.36	0.38	-0.183	0.35	0.35	0.001
Baseline literacy	0.51	0.58	-0.319	0.51	0.51	-0.032
Baseline sanitation	0.96	0.93	0.228	0.98	0.98	-0.023
Baseline pop. density	3.5	20.6	-1.892	3.7	4.0	-0.023
Travel time to pop. centre	4461	1956	0.943	4199	3889	0.116

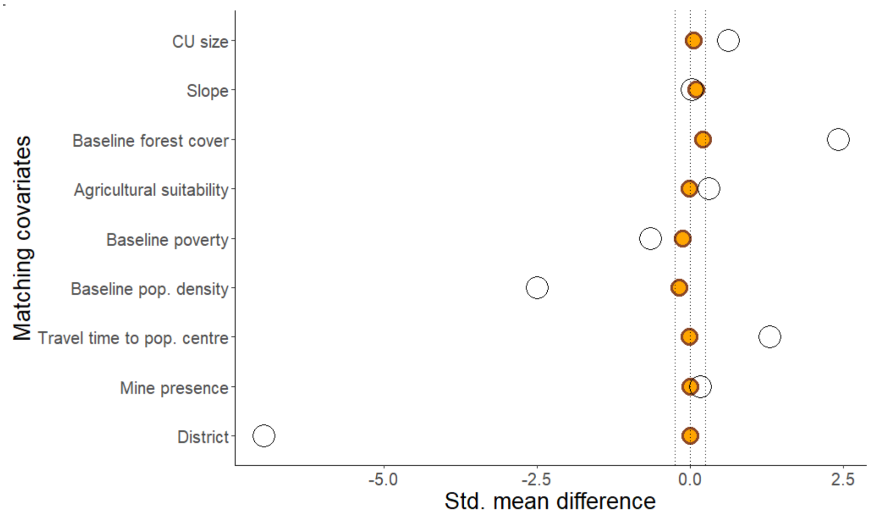

Figure S3.9. Balance of covariates before and after genetic matching, using sustainable use protected areas established after 2000 as treatment, very small landholders as control and deforestation as outcome variable. Standardized mean difference for confounding covariates before (open circles) and after matching (orange circles). Dotted lines indicate acceptable standard mean difference after matching to reach balance (Stuart 2010).

Table S3.6. Balance of covariates before and after genetic matching, using sustainable use protected areas established after 2000 as treatment, non-protected areas as control and poverty indicators as outcome variables. Analysis was repeated for each poverty indicator separately, but only covariate balance for income outcome is reported. For other poverty indicators only post-matched balance of the poverty indicator is reported.

	Before matching			After matching		
	Means treated (n = 176)	Means control (n = 438)	Std. mean Diff.	Means treated (n = 75)	Means control (n = 28)	Std. mean Diff.
CT size	3084	306	0.617	677	472	0.046
Slope	29.9	26.7	0.027	63.6	60.7	0.024
Baseline forest cover	91.5	60.7	2.415	87.1	84.3	0.223
Agricultural suitability	0.8	0.6	0.307	1.06	1.06	-0.017
Baseline income	386	573	-0.451	350	403	-0.128
Baseline GINI	0.36	0.40	-0.304	0.35	0.34	0.074
Baseline literacy	0.51	0.61	-0.474	0.53	0.54	-0.039
Baseline sanitation	0.96	0.90	0.509	0.97	0.98	-0.070
Baseline pop. density	3.5	26.0	-2.497	7.3	8.6	-0.141
Travel time to pop. centre	4461	1018	1.297	3062	3098	-0.014

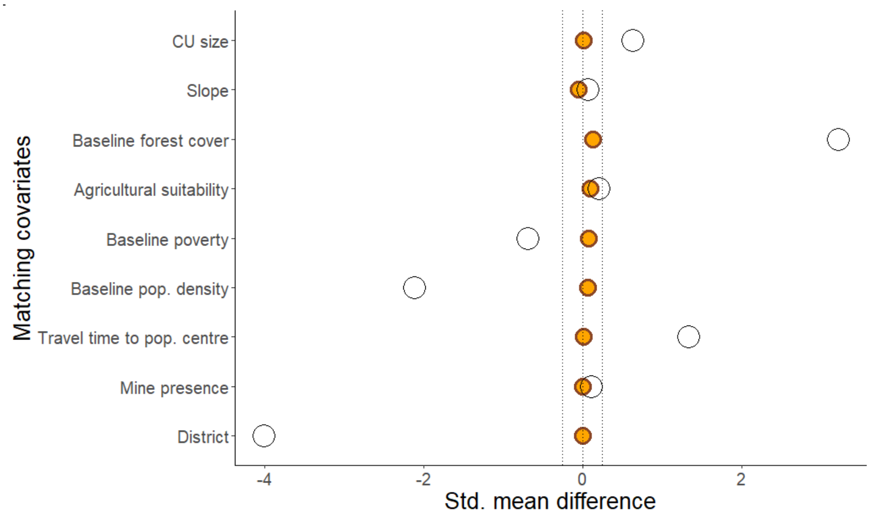

Figure S3.10. Balance of covariates before and after genetic matching, using sustainable use protected areas established after 2000 as treatment, smallholders as control and deforestation as outcome variable. Standardized mean difference for confounding covariates before (open circles) and after matching (orange circles). Dotted lines indicate acceptable standard mean difference after matching to reach balance (Stuart 2010).

Table S3.7. Balance of covariates before and after genetic matching, using sustainable use protected areas established after 2000 as treatment, non-protected areas as control and poverty indicators as outcome variables. Analysis was repeated for each poverty indicator separately, but only covariate balance for income outcome is reported. For other poverty indicators only post-matched balance of the poverty indicator is reported.

	Before matching			After matching		
	Means treated (n = 176)	Means control (n = 1084)	Std. mean Diff.	Means treated (n = 45)	Means control (n = 26)	Std. mean Diff.
CT size	3084	273	0.624	529	501	0.006
Slope	29.9	3.6	0.222	17.1	11.2	0.050
Baseline forest cover	91.5	52.0	3.098	87.0	86.1	0.075
Agricultural suitability	0.80	0.59	0.402	0.91	0.94	-0.055
Baseline income	386	602	-0.523	334	315	0.048
Baseline GINI	0.36	0.45	-0.614	0.38	0.38	0.023
Baseline literacy	0.51	0.65	-0.652	0.59	0.57	0.132
Baseline sanitation	0.96	0.95	0.096	0.96	0.97	-0.038
Baseline pop. density	3.5	15.5	-1.327	6.6	8.4	-0.200
Travel time to pop. centre	4461	993	1.306	2502	2564	-0.023

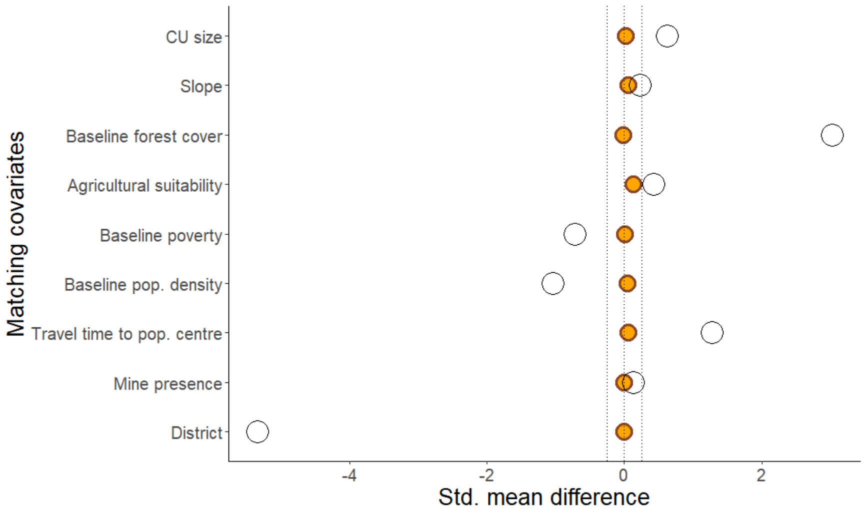

Figure S3.11. Balance of covariates before and after genetic matching, using sustainable use protected areas established after 2000 as treatment, medium-sized landholders as control and deforestation as outcome variable. Standardized mean difference for confounding covariates before (open circles) and after matching (orange circles). Dotted lines indicate acceptable standard mean difference after matching to reach balance (Stuart 2010).

Table S3.8. Balance of covariates before and after genetic matching, using sustainable use protected areas established after 2000 as treatment, non-protected areas as control and poverty indicators as outcome variables. Analysis was repeated for each poverty indicator separately, but only covariate balance for income outcome is reported. For other poverty indicators only post-matched balance of the poverty indicator is reported.

	Before matching			After matching		
	Means treated (n = 176)	Means control (n = 1547)	Std. mean Diff.	Means treated (n = 75)	Means control (n = 41)	Std. mean Diff.
CT size	3084	290	0.620	786	684	0.023
Slope	29.9	3.3	0.224	17.9	10.6	0.061
Baseline forest cover	91.5	52.8	3.032	88.1	88.3	-0.016
Agricultural suitability	0.8	0.58	0.426	0.89	0.85	0.090
Baseline income	386	617	-0.560	329	304	0.061
Baseline GINI	0.36	0.44	-0.575	0.344	0.336	0.061
Baseline literacy	0.51	0.66	-0.697	0.53	0.54	-0.072
Baseline sanitation	0.96	0.95	0.075	0.98	0.98	-0.013
Baseline pop. density	3.5	13.0	-1.048	4.7	3.6	0.127
Travel time to pop. centre	4461	1085	1.272	3157	3036	0.046

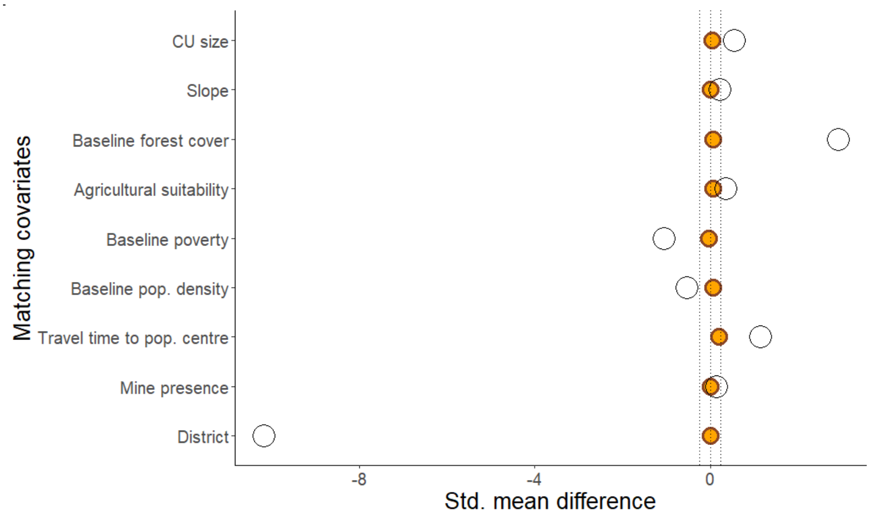

Figure S3.12. Balance of covariates before and after genetic matching, using PAs established after 2000 as treatment, large landholders as control and deforestation as outcome variable. Standardized mean difference for confounding covariates before (open circles) and after matching (orange circles). Dotted lines indicate acceptable standard mean difference after matching to reach balance (Stuart 2010).

Table S3.9. Balance of covariates before and after genetic matching, using sustainable use protected areas established after 2000 as treatment, non-protected areas as control and poverty indicators as outcome variables. Analysis was repeated for each poverty indicator separately, but only covariate balance for income outcome is reported. For other poverty indicators only post-matched balance of the poverty indicator is reported.

	Before matching			After matching		
	Means treated (n = 176)	Means control (n = 582)	Std. mean Diff.	Means treated (n = 71)	Means control (n = 28)	Std. mean Diff.
CT size	3084	593	0.553	850	780	0.016
Slope	29.9	3.9	0.220	21.1	21.7	-0.006
Baseline forest cover	91.5	54.1	2.929	90.4	89.3	0.087
Agricultural suitability	0.80	0.62	0.349	0.93	0.95	-0.033
Baseline income	386	889	-1.217	353	397	-0.106
Baseline GINI	0.36	0.46	-0.735	0.36	0.39	-0.227
Baseline literacy	0.51	0.70	-0.883	0.50	0.51	-0.055
Baseline sanitation	0.96	0.92	0.309	0.97	0.98	-0.043
Baseline pop. density	3.5	8.3	-0.534	4.4	2.8	0.169
Travel time to pop. centre	4461	1391	1.156	3259	2788	0.177

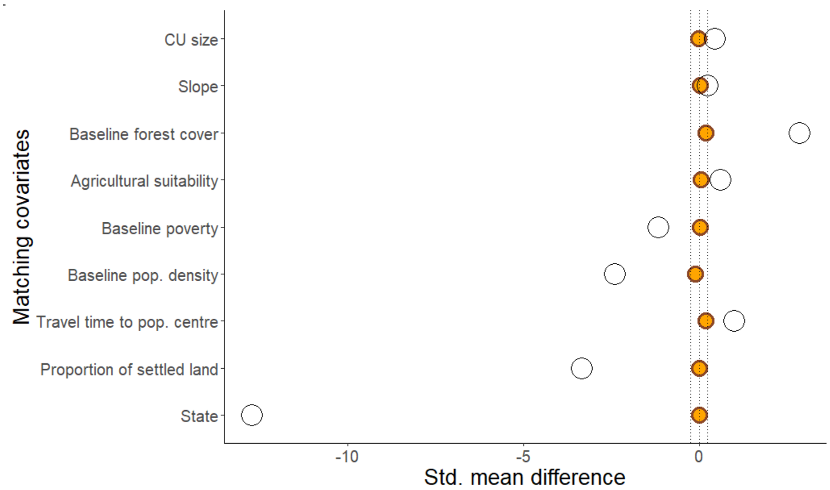

Figure S3.13. Balance of covariates before and after genetic matching, using PAs established after 2000 as treatment, mines established after 2000 as control and deforestation as outcome variable. Standardized mean difference for confounding covariates before (open circles) and after matching (orange circles). Dotted lines indicate acceptable standard mean difference after matching to reach balance (Stuart 2010).

Table S3.10.| Balance of covariates before and after genetic matching, using sustainable use protected areas established after 2000 as treatment, non-protected areas as control and poverty indicators as outcome variables. Analysis was repeated for each poverty indicator separately, but only covariate balance for income outcome is reported. For other poverty indicators only post-matched balance of the poverty indicator is reported.

	Before matching			After matching		
	Means treated	Means control	Std. mean Diff.	Means treated	Means control	Std. mean Diff.
CT size	2309	481	0.444	1275	1215	0.015
Slope	31.3	2.4	0.244	11.9	7.7	0.036
Baseline forest cover	89.5	49.6	2.862	88.0	85.4	0.189
Agricultural suitability	0.86	0.56	0.617	0.81	0.81	0.036
Baseline income	347	782	-2.049	347	334	0.060
Baseline GINI	0.36	0.46	-0.784	0.36	0.36	-0.065
Baseline literacy	0.51	0.70	-0.925	0.56	0.55	0.046
Baseline sanitation	0.96	0.90	0.568	0.98	0.97	0.077
Baseline pop. density	5.3	37.4	-2.403	4.6	5.0	-0.029
Travel time to pop. centre	3954	1201	1.002	2854	2194	0.240
Proportion of settled land	0.07	0.50	-3.329	0.08	0.09	-0.083

References

Ali, S.H., Giurco, D., Arndt, N., Nickless, E., Brown, G., Demetriades, A., Durrheim, R., Enriquez, M.A., Kinnaird, J., Littleboy, A., Meinert, L.D., Oberhänsli, R., Salem, J., Schodde, R., Schneider, G., Vidal, O. & Yakovleva, N. (2017). Mineral supply for sustainable development requires resource governance. *Nature*, 543, 367–372.

Andam, K.S., Ferraro, P.J., Sims, K.R.E., Healy, A. & Holland, M.B. (2010). Protected areas reduced poverty in Costa Rica and Thailand. *Proc. Natl. Acad. Sci. U. S. A.*, 107, 9996–10001.

Angelsen, A. (2010). Policies for reduced deforestation and their impact on agricultural production. *Proc. Natl. Acad. Sci. U. S. A.*, 107, 19639–19644.

Canavire-Bacarreza, G. & Hanauer, M.M. (2013). Estimating the Impacts of Bolivia's Protected Areas on Poverty. *World Dev.*, 41, 265–285.

DNPM, D.N. de P.M. (2012). Processos Minerarios: Sistema de Informações Geográficas da Mineração (SIGMINE) [WWW Document]. URL http://sigmine.dnpm.gov.br/

Fischer, G.F., Nachtergaele, S., Prieler, H.T., van Velthuizen, H.., Verelst, L. & Wiberg, D. (2007). *Global Agro-ecological Zones Assessment for Agriculture*.

Gentle, P. & Maraseni, T.N. (2012). Climate change, poverty and livelihoods: adaptation practices by rural mountain communities in Nepal. *Environ. Sci. Policy*, 21, 24–34.

Hecht, S.B. (1985). Environment, development and politics: Capital accumulation and the livestock sector in Eastern Amazonia. *World Dev.*, 13, 663–684.

Joppa, L.N. & Pfaff, A. (2009). High and Far: Biases in the Location of Protected Areas. *PLoS One*, 4.

Nepstad, D.C., Stickler, C.M., Filho, B.S.- & Merry, F. (2008). Interactions among Amazon land use, forests and climate: prospects for a near-term forest tipping point. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.*, 363, 1737–46.

Nolte, C., Agrawal, A., Silvius, K.M. & Soares-Filho, B.S. (2013). Governance regime and location influence avoided deforestation success of protected areas in the Brazilian Amazon. *Proc. Natl. Acad. Sci. U. S. A.*, 110, 4956–4961.

Soares-Filho, B.S., Nepstad, D.C., Curran, L.M., Cerqueira, G.C., Garcia, R.A., Ramos, C.A., Voll, E., McDonald, A., Lefebvre, P. & Schlesinger, P. (2006). Modelling conservation in the

Amazon basin. *Nature*, 440, 520–523.

Sonter, L.J., Ali, S.H. & Watson, J.E.M. (2018). Mining and biodiversity: key issues and research needs in conservation science. *Proc. R. Soc. B Biol. Sci.*, 285, 20181926.

Stuart, E.A. (2010). Matching Methods for Causal Inference: A Review and a Look Forward. *Stat. Sci.*, 25, 1–21.

Wittemyer, G., Elsen, P., Bean, W.T., Burton, A.C.O. & Brashares, J.S. (2008). Accelerated human population growth at protected area edges. *Science (80-.).*, 321, 123–126.

Chapter 4:

Testing the Ability of Machine Learning to Estimate Poverty and Poverty Change in Tanzania Using Publicly Available Data Sources

Bowy den Braber, Gary Watmough, Maurico Alvarez, Karl L. Evans, Johan A. Oldekop

B.d.B., J.A.O. and K.L.E. designed research; B.d.B. performed research and analysed data; G.W. and M.A. provided technical input; B.d.B. wrote the paper with input from all authors

Met opmerkingen [b14]: Auhtor contribution

4.1 Abstract

Reducing poverty is a key target of sustainable development goals, which will require monitoring poverty at multiple time points and at fine spatial resolutions. Poverty data in lower- and middle-income countries are often insufficiently frequent or at too coarse spatial resolutions to use for assessing the impacts of environmental change or poverty reduction policies. Previous studies have been able to predict poverty fairly accurately using machine learning methods (convolutional neural networks (CNNs)) and very high resolution satellite imagery. Here we assess the ability of CNNs to estimate poverty and poverty change in Tanzania using publicly available data sources. We find that CNNs provide useful information to estimate household consumption expenditure (Pearson's r^2 0.32), but cannot be used to measure poverty change or multidimensional poverty. Estimates vary per biome and are less useful to detect small changes in poverty. Our results can be used to improve poverty predictions in countries with limited socio-economic data.

4.2 Introduction

Eradicating extreme poverty by 2030 is a key target of the sustainable development goals (SDGs - Griggs et al. 2013). Monitoring progress towards this target, and understanding how social and environmental factors, and policy change influence poverty dynamics requires quantitative assessments of poverty at multiple time points and at sufficiently fine spatial resolutions to capture spatial heterogeneity (Lu et al. 2015; Selomane et al. 2015). These data are crucial for policy makers and NGOs to target areas and people most in need of support, and to evaluate the effectiveness of poverty reduction efforts. Most programs and policies operate over short time frames, and their implementation is often spatially heterogeneous. Yet, the majority of poverty data in lower- and middle-income countries is gathered through census cycles or representative household surveys (e.g., the World Bank's Living Standards Measurement Survey) that collect data intermittently over space and time, and are thus not suitable to assess the effects of dedicated programs and policies.

Sub-Saharan Africa is the region with the largest data gap. Rapid economic and demographic changes in the region are not matched by suitable socio-economic data, which can result in poor quality data and misleading official statistics that do not represent the reality on the ground (Jerven 2013). Consequently, the United Nations High Level Panel has called for a data revolution to dramatically improve data collection (United Nations IEAG 2014). However, collecting such data requires expensive and time-consuming household surveys that continue to be inconsistently available over time and space. Following guidelines set up by the World Bank for monitoring the SDGs, such a data collection revolution is estimated to cost around $253 billion globally for the time period of the SDGs – almost twice the annual global total spent on official development assistance (Jerven 2014).

An alternative, and potentially much cheaper, approach to reducing the data gap is to use secondary data sources to predict poverty at fine spatial scales across large spatial extents. This general approach has been applied using mobile phone network data (Blumenstock et al. 2015; Steele et al. 2017), and remote sensed satellite data (Elvidge et al. 2009) – with the later showing the most promise due to its global availability and scalability in comparison to other data sources. These advantages will continue to increase as the spatial and temporal resolution and availability of satellite images continues to improve - for example through the launching of the Sentinel satellites (Drusch et al. 2012). Data derived from satellite sensors are used to extract information on land use and other biophysical characteristics, but can also be used as proxies to measure socio-economic development and livelihoods. Luminosity (nightlight) measured by satellites has been shown to be a good proxy for economic activity (Noor et al. 2008; Chen & Nordhaus 2011). For example, Noor et al. (2008) showed that nightlight correlated with an asset-based poverty index at 64% precision. However, in areas with high poverty, luminosity levels are generally very low or close to zero making it difficult to estimate small changes in poverty using night-time light data in isolation.

Met opmerkingen [b15]: Example to clarify point

Recently, studies have applied novel statistical methods using remotely sensed data from satellites to estimate poverty, education, and health indicators in sub-Saharan Africa (Golding et al. 2017; Graetz et al. 2018; Osgood-Zimmerman et al. 2018; Tusting et al. 2019). Machine learning methods have shown promise with regard to their ability to predict poverty using information derived from remotely sensed imagery (Watmough et al. 2016, 2019). For instance, Watmough et al. (2019) estimated poverty using very high resolution imagery and household surveys up to 62% accuracy. One approach used by Jean et al. (2016) applies convolutional neural networks (CNNs) and imagery from Google Street Maps to measure poverty in four African countries at higher accuracy than using luminosity. CNNs are a type of deep neural networks that are developed specifically for image classification tasks (Lecun et al. 1998; Bengio 2009). In contrast to other machine learning methods they are able to derive information from multiple pixels simultaneously, and are thus able to detect spatial features from images. Jean et al. (2016) show that CNNs are able to detect landscape features from imagery that correlate with economic activity and are thus useful for detecting poverty. New forms of CNNs have been developed to incorporate a temporal dimension to classify change in imagery. These so-called 3-dimensional CNNs are originally developed to classify video footage (Karpathy et al. 2014), but have been successful in detecting changes in imagery in various research domains (Ji et al. 2013, 2018).

Met opmerkingen [b16]: Example to clarify point

Several topics remain understudied. First, current studies that apply machine learning methods to estimate poverty typically use very high-resolution imagery (~3m/pixel) that is expensive, not publicly available, and collected irregularly over space and time because observations are captured 'on demand' (Watmough et al. 2019), but see Perez et al. (2019). Although these high-resolution images are better for detecting features from satellite imagery (Gilbertson et al. 2017), the costs of these images hinder standardized and routine poverty data estimations. In contrast, Landsat imagery has a lower resolution (30m), but is publicly available from 1984 and continues to be collected regularly (16-day repeat cycle). Studies to date have not assessed if these advantages of using Landsat imagery trade off against reduced capacity to predict poverty due to the lower resolution imagery. A key challenge is therefore to assess the ability to measure poverty using lower resolution publicly available imagery.

Second, currently the state-of-the-art is single time point poverty maps that only provide snapshots of poverty (cross-sectional data), but do not provide insight into the dynamics of poverty over time (longitudinal data). However, accurately measuring change is crucial to monitor progress towards the SDGs and to conduct impact evaluations of specific policies or programs. Single time-point maps are not suitable to assess changes over time, because measurement errors of individual time points are compounded when combined to measure change. Compounding errors are commonly addressed in deforestation change measures (see e.g. Sexton et al. 2015), but are currently understudied for poverty change measures. A second key challenge is therefore to develop methods to estimate remotely-sensed poverty change that minimize errors.

Finally, previous studies have mostly focused on consumption-based metrics. Consumption- or income-based metrics are commonly used to measure poverty, but are also criticized as poor indicators of people's needs (Anand & Sen 1997). Other dimensions of human welfare such as education, sanitation and assets are important aspects in driving wellbeing and do not always correlate well with consumption or income but are recommended for use in multi-dimensional poverty indices (Alkire & Foster 2011). The final key challenge is thus to assess the capacity of machine learning methods applied to remote sensing data for accurately measuring poverty dimensions other than consumption.

Here, we assess the ability to predict poverty change at fine spatial scales in Tanzania using publicly available Landsat satellite imagery and novel CNN methods. We obtain panel household consumption expenditure data (a key indicator of poverty) from the World Bank's Living Standards Measurements Study – Integrated Surveys on Agriculture (LSMS-ISA) from 2008/09 and 2012/13 to estimate poverty for each year and directly estimate poverty change from 2008/09 to 2012/13. In addition, we estimate multi-dimensional poverty using data from the Demographic Health Survey (DHS) in 2010. We apply

CNN transfer learning models that have shown to be successful in estimating poverty indicators (Jean et al. 2016; Xie et al. 2016), and apply specific 3D CNNs to predict poverty change. We estimate consumption-based poverty and multidimensional poverty in the two biggest biomes (tropical forest and tropical grassland biomes) and compare the accuracy of our poverty estimates. We use Tanzania as a case study because it has major gaps in coverage from censuses, contains major variation in biomes, has very high poverty rates, and evidence for marked recent changes in poverty levels (Ellis & Mdoe 2003; Östberg et al. 2018).

4.3 Methods

4.3.1 General approach

We used CNNs to estimate household consumption expenditure, an important indicator of poverty, in Tanzania in 2008/09, 2012/13 and estimate consumption change from 2008/09 to 2012/13. In addition, we estimate multidimensional poverty in 2010. Following Jean et al. (2016) we used a transfer learning approach consisting of three steps (Figure 4.1). First, we extracted day-time (Landsat 5 TM and Landsat 8 OLI/TIRS) and night-time imagery from publicly available satellite data. For every location we classified day-time imagery onto a corresponding nightlight intensity category (binned into categories from low to high luminosity levels). Secondly, we split our imagery into training and validation sets (70/30 ratio) and trained a convolutional neural network to identify features (e.g. roads, buildings or abstract features) that are useful to classify day-time imagery into a corresponding nightlight category. Features from day-time images that are useful to predict nightlights can also be useful to predict poverty levels (Jean et al. (2016)). Finally, we use the extracted features from the CNN for further analyses to estimate how much of the poverty can be explained by the extracted features. Ideally, the imagery used for the nightlights categorization task should be separate from the imagery used to estimate poverty. Few poverty data points restrict the possibility to split the data in training and validation data (for estimating nightlights), and testing data (for estimating poverty). Moreover, the dense sampling of the LSMS-ISA and DHS surveys in Tanzania in villages and urban areas makes it impossible to select separate imagery for categorizing nightlights without overlap (Figure S4.1). Other studies have overcome this issue by sampling imagery for the nightlight categorization task from other countries (Jean et al. 2016; Xie et al. 2016), but this is likely to introduce error due to differences between countries in environmental, socio-economic and cultural factors that generate international variation in the relationship between poverty, nightlights and features extracted from remote sensed data. Since our main task is not to classify nightlights but estimate poverty we decided to use the same imagery for the nightlight categorization task to extract features for estimating poverty, although the same network structure could be used to handle a much larger set of images which would likely improve predictive capacity of poverty. In addition, we used an existing well-established CNN architecture that is not trained on nightlight imagery, but uses weights pre-trained on a standardized image database (ImageNet) to extract features and compare our results.

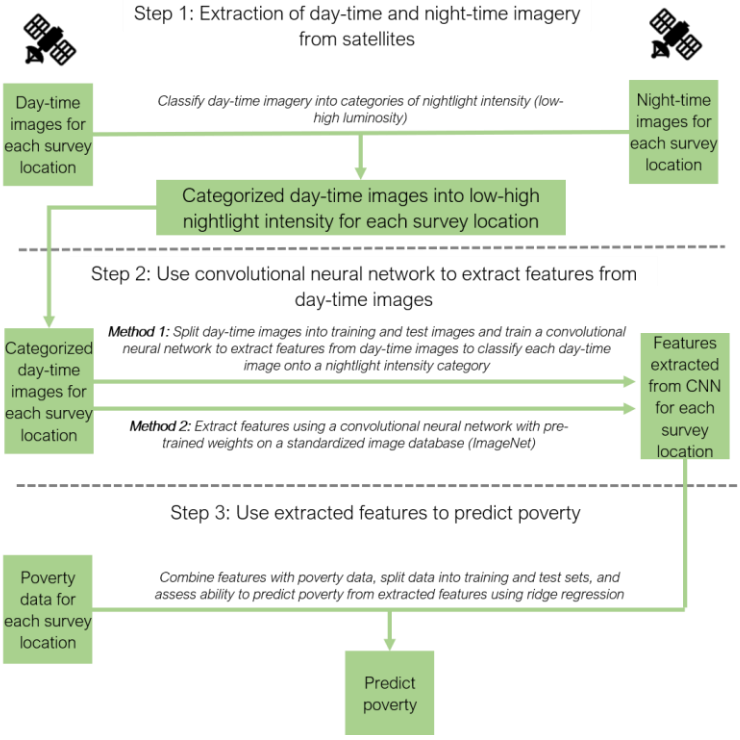

Figure 4.1. Schematic overview of general approach to predict poverty using satellite imagery. In step 1 we extract (10x10 km^2) day-time and night-time images from satellite data at each survey location from the household surveys (Living Standards Measurements Study – Integrated Surveys on Agriculture (LSMS-ISA) and Demographic Health Survey (DHS)). Then for every location we overlay day-time and night-time imagery and classify every day-time image onto a corresponding nightlight intensity category (ranging from low to high luminosity levels). In step 2 we split our day-time images into training and validation images (70/30 ratio) and train a convolutional neural network (CNN) to extract features (e.g. roads, buildings or abstract features) that are useful to classify day-time imagery onto a corresponding nightlight category. In step 3 we use the extracted features from the CNN and combine these with the poverty data. We split the data into training and test sets (70/30) and assess the ability to predict poverty from the extracted features using ridge regression.

4.3.2 Study site

Tanzania experienced significant economic and demographic growth since the 1990s (Ellis & Mdoe 2003; Östberg et al. 2018) as a result of the booming mining, tourism, financial and construction sectors. Agricultural growth has been stagnant, yet 68% of the population continues to live in rural areas, including 84% of the country's poor. Poverty rates as measured by the LSMS-ISA surveys have declined since 1992, but because of population growth the number of poor people in Tanzania has increased over the same period. Yet much remains unknown about local poverty dynamics (Östberg et al. 2018). Improved poverty data would aid in determining effects of multiple policies and programs that affect human wellbeing (Pailler et al. 2015).

4.3.3 Poverty data

We used two different data sources to measure poverty in Tanzania that enable us to measure consumption based poverty and multi-dimensional poverty. First, we extracted poverty household consumption expenditure data in 2008/2009 and 2012/2013 from the LSMS-ISA that are collected by the World Bank (Grosh & Glewwe 1998). Consumption expenditure is a common measure of poverty and is defined as the total amount of money a household spends on consumption goods per year. LSMS-ISA surveys in Tanzania and other African countries are part of a longitudinal study to track household poverty over time and are thus suitable to estimate poverty change. LSMS-ISA surveys sample enumeration areas ("clusters") throughout the country with probability of sampling in proportion to population size. Households within the cluster are then randomly sampled. LSMS-ISA report the latitude/longitude of the cluster, but geolocations are made fuzzy to ensure anonymity of individuals is preserved. Up to 2 km noise is added for urban areas, 5 km for rural areas, and 10 km is randomly added for 1% of the cluster. We extracted imagery from a 10 by 10 km square polygon around the cluster geolocation. This way we take into account the fuzzy nature of the geolocation, but also capture a wider area around the household which is more representative of their economic well-being as economic activities of a household take place in a buffer around the household (Watmough et al. 2019).

Second, we constructed a multidimensional poverty index using data obtained from the Demographic Health Surveys (DHS). Although consumption poverty is commonly used as a poverty metric, certain aspects of poverty are overlooked (Brockington et al. 2018). From the DHS data we constructed a multidimensional poverty index, based on education, health and living standards (Table S4.1 - Alkire

& Foster 2011). We obtained data from 2010, the closest year to our LSMS-ISA data. Sampling design is similar to the LSMS and locations of clusters are also made fuzzy. However, the DHS data sampling design is not longitudinal, and is therefore not suitable to track poverty over time at fine spatial resolutions and therefore we only predicted multidimensional poverty in 2010. Similarly, we extracted an image for every 10 by 10 km square polygon drawn around the sampled geolocation.

4.3.4 Day-time Landsat satellite imagery

We obtained cloud-free composite Landsat satellite imagery from Tanzania in 2008 and 2013, corresponding to the years of the available poverty data. The resolution of Landsat imagery (30m/pixel for most bands, 15m/pixel for pan-sharpened band) is significantly lower than the very high resolution imagery (~2.5m/pixel) that has previously been used to assess poverty (Jean et al. 2016; Watmough et al. 2019), but Landsat imagery is open source, easily accessible and has been collected since 1984 (Wulder et al. 2016). All acquiring and processing of imagery was done in Google Earth Engine (Gorelick et al. 2017). We used atmospherically corrected Landsat 5 TM Tier 1 images for 2008 and Landsat 8 OLI/TIRS Tier 1 images for 2013. Atmospherically corrected Tier 1 products provide similar spectral signatures and give the highest geodetic accuracy and are thus recommended for time series analyses (Young et al. 2017). We chose to avoid using Landsat 7 because of the scan-line corrector failure that would have unpredictable effects on our CNN models. Because our analysis draws from two different sensors we cross-calibrated Landsat 8 surface reflectance values to Landsat 5 reflectance to ensure cross-sensor functionality (Roy et al. 2016). Although we use the algorithm to cross-calibrate Landsat 8 surface reflectance to Landsat 7, reflectance values of the Landsat 5 TM and Landsat 7 ETM+ sensors are very similar (Melaas et al. 2013). For both years we extracted the visual bands (red-green-blue (RGB)), near-infrared (NIR), both shortwave infrared bands (SWIR), and the thermal infrared (TIR) band for our analyses. We estimated models that only use the visual RGB bands, and models that use all seven bands. For 2013 we also tested Landsat 8 OLI/TIRS Tier 1 top-of-atmosphere (TOA) imagery because it allows to pan-sharpen the Red-Green-Blue bands to a higher resolution (15m), which has proven to be successful in other image classification studies (Gilbertson et al. 2017). Pan-sharpened imagery is not available for Landsat 5 TM so could only be used for 2013 data. We selected the images with the fewest clouds using the CFMASK algorithm (Foga et al. 2017) and produced a composite for every year by taking the median value of each pixel from that year during the period June-October, which is considered the dry period for most of Tanzania. Finally, we overlaid our cluster polygons from the LSMS-ISA and DHS datasets with the final composite map to extract day-time images. In total we ended up with 369 images for 2008, 1458 images for 2013, and 409 images for 2010 (DHS). To estimate change we used the 369 images for 2008 and used the same geolocations to extract 369 images for 2013.

4.3.5 Nightlight satellite imagery

We obtained night-time satellite imagery in 2008 and 2013 from the United States Air Force Defense Meteorological Satellite Program (DMSP), provided by the National Oceanic and Atmospheric Administration (NOAA). Although the more recent Visible Infrared Imaging Radiometer Suite (VIIRS) data provides a higher resolution (Elvidge et al. 2013), this data is not available for 2008 and we therefore decided to only use the DMSP data in our analyses. Previous work on estimating poverty has also used the DMSP data (Jean et al. 2016). After extracting the nightlight data we overlaid our LSMS-ISA and DHS cluster polygons with the nightlight data to determine the average nightlight luminosity for each cluster polygon. For classification purposes we then binned the nightlight data into 5 categories ranging from no light to very bright so that every day-time image corresponds to a nightlight category. For our model to estimate poverty change we binned the change in luminosity between the years in 3 categories (no change, small change and big change).

4.3.6 Convolutional neural networks

We used CNNs to extract features from the day-time imagery that can be used to predict poverty. Our CNN models were based on the VGG-16 model structure (Simonyan & Zisserman 2014) which performs well on classifying images in the widely used ImageNet database (Russakovsky et al. 2015). We used Keras (Chollet 2015), based on Tensorflow (Abadi et al. 2016), to implement our CNN models. The VGG-16 model architecture consists of convolutional layers, pooling layers, and a fully connected layer (Figure S4.2). In the convolutional layers features are extracted from an image while keeping the spatial dependency. Relatively small (3x3 pixels) filters are passed over an image to create generalized feature maps. The number of filters increases per block of convolutional layers from 32 filters in the first block to learn basic features such as edges to 512 filters in the last block to learn complex features such as objects. We added non-linearity in the form of a rectified linear unit (ReLU) to allow complex relationships to be learned (Dahl et al. 2013). Batches of convolutional layers are followed by pooling layers. Pooling layers are used to reduce dimensionality and prevent overfitting by minimizing the number of parameters in the model. In the final step of the network a fully connected layer combines the neuron information from the previous layer to determine the night light category of the image. We then applied the SoftMax activation function that is commonly used for image classification purposes (Bengio 2009). In our transfer learning models we first train the images to classify day-time images to a corresponding nightlight intensity category and then extract the features from the last pooling layer before the fully connected layer for further analyses. For all models we used dropouts and data augmentation (horizontal and vertical flipping, and rotating of images by 90 degrees) to prevent overfitting (Srivastava et al. 2014). We used categorical cross entropy as our loss function, an Adam optimizer starting at 0.0001 (Kingma & Ba 2014), a batch size of 32, and input image resolution of 128 by 128 as other CNN parameters.

To estimate poverty change we altered the original network to allow two images as input and create a 3D CNN (Figure S4.3). 3D CNNs have originally been created to analyse video footage, but can also be used to assess temporal change (Donahue et al. 2015). In this case, the CNN receives two day-time images, one from 2008 and one from 2013, and assesses the change between the images directly and classifies every pair of images to a corresponding nightlight intensity change category. Poverty change is thus modelled directly from changes in the images rather than estimating poverty separately for two different time periods and thus avoids compounding errors. Just like in our transfer learning CNN we extracted the features from the last pooling layer before the fully connected layer for further analyses.

One common approach is to apply a VGG-16 CNN that is already pre-trained on ImageNet and use this CNN as a feature extractor. These features can then be used in a different task. This approach has shown to be useful in a variety of fields (Pan & Yang 2010; LeCun et al. 2015). Using pre-trained weights is useful to detect low-level features such as edges and slows down computational time (Krizhevsky et al. 2012). However, ImageNet mainly consists of object-based images whereas satellite imagery is observed from an aerial, bird's-eye view (Christie et al. 2018). However, extracting features from CNNs trained on ImageNet has also shown to be useful in remote sensing applications (Nogueira et al. 2017; Hu et al. 2015).

We tested both models that use the VGG-16 trained on ImageNet as a feature extractor and transfer models that are trained from scratch to classify nightlights and train all layers of the CNN. Because VGG-16 trained on ImageNet only accepts three bands we only used the pan-sharpened images and the normal RGB images for this approach. We used all different imagery (pan-sharpened, normal RGB and all bands) for our transfer learning models that are trained from scratch on nightlights.

4.3.7 Post-CNN

The features that are extracted from CNNs can be further used in regression analyses to estimate household consumption and multidimensional poverty metrics. Our transfer learning model CNN and the ImageNet VGG-16 CNN extract 5048-dimensional feature vectors (our change models extract

10096-dimensional feature vectors). To reduce overfitting and to enhance computational speed we first used principal component analysis to reduce dimensionality (Jean et al. 2016). We used the first 50 components of the PCA (Jean et al. 2016) and ran ridge regressions which further prevents overfitting by penalizing the size of the linear coefficients (R package 'glmnet'- Friedman et al. 2009). We computed the average Pearson's R^2 correlation after nested 5-folded cross-validation. In the inner fold we tune the alpha L2 regularization parameter. In each outer fold we split our poverty data in training (70%) and test (30%) data. We repeat this procedure to calculate the capacity to predict poverty per biome (tropical forest and tropical grassland biomes), and to predict poverty while using different poverty thresholds (using only 25%-50%-75% of the poorest clusters). The latter tests whether CNNs can also detect small-scale heterogeneity. To assess poverty change we did not use different poverty thresholds, but instead used different change thresholds (using 25%-50%-75% of the clusters with smallest change). For every approach we compared our correlations using CNNs against the correlation between nightlight and the poverty metric. We logged consumption expenditure so that consumption was distributed normally.

Secondly, we divided all our continuous poverty metrics in four different bins to check the robustness of our estimates and evaluate the impacts of false classifications. We used a support vector machine (SVM) to predict which category our test poverty data belongs to. SVM is a supervised machine learning algorithm that separates classes using a multidimensional non-linear approach (Kuhn 2008). Again we used nested 5-fold cross-validation. In the outer fold the poverty data is split into training (70%) and test (30%) data while in the inner fold the SVM parameters are fine-tuned during each fold. We used two different metrics to assess the classification accuracy. The first is the multiclass AUC which is the area under the receiver operating characteristic curve (ROC). The ROC curve compares the true positive rate against the true negative rate. The AUC is 0.5 when classification is completely random and 1 indicates perfect prediction. Secondly, we calculate weighted Cohen's kappa because, other than AUC, it respects the ordinal structure of the poverty classes that ranges from low to high. Weighted Cohen's kappa extra penalizes classifications that are two or three classes away from the true category. Because nightlight is a continuous variable we only calculated the AUC in this case.

4.4 Results

We show that CNNs trained on publicly available satellite imagery can be used to predict average household consumption expenditure in 2008/09 and 2012/13. Cross-validated predictions can explain 31% of the variation in 2008/09 consumption expenditure and 32% of the variation in 2012/13 (Figure 4.2). These estimates are lower than the estimates found using very high-resolution imagery (55% in

Tanzania by Jean et al. (2016)). They are also lower than the estimates when using nightlight intensity as a predictor of consumption. Nightlight luminosity can explain over 50% of the variation both for 2008/09 and 2012/13 (Figure 4.2A-B).

Our ability to assess change over time is lower than our predictions for single time snapshots. Still our best model can explain 14% of the variation despite the relatively short time span of only four years (Figure 4.2 C). However, changes in luminosity predict changes in consumption slightly better (17%).

We also find lower predictive capacity to estimate multidimensional poverty compared to our consumption-based models. Our CNNs can only explain 16% of the variation in the MPI in 2010 (Figure 4.2D). This could be explained by the nature of the index that also includes dimensions of poverty that are difficult to assess using satellite imagery such as education. Nightlights can predict multidimensional poverty slightly better (29%) but also show lower predictive power in comparison to household consumption predictions.

Predictive capacity highly depends on the focal biome (Figure 4.2). For all poverty measures predictive power in the tropical forest biome that spans the coastal area is much higher than in the tropical grasslands that span much of the rural part of Tanzania. In the latter predictive capacity of our CNN models drops to almost zero. Nightlight performs slightly better in that part of Tanzania, but also shows a drop in explanatory capacity. In some cases predictive capacity in the tropical forest biome is higher than across Tanzania. Our best CNN is able to explain 53% of the variation in household consumption in this biome (compared to 32% for entire Tanzania), even though this is still lower than nightlight predictions (69%). We find similar patterns when assessing poverty change. However, we also find that one model predicted poverty change in the tropical forest biome better than using nightlight as a predictor of consumption (29% cf. 21%).

Our models cannot detect heterogeneity in consumption and multidimensional poverty among poorer household clusters. Predictive capacity of all tested models drops to almost zero when only the poorest quantiles of the household clusters are measured. Nightlight intensity performs equally poor to measure heterogeneity in poverty close to the poverty line (Figure 4.3). Our poverty change models also show drops in predictive capacity when assessing heterogeneity in smaller changes, although one CNN model shows a sudden increase in predictive capacity to assess small changes (18%- Figure 4.3C).

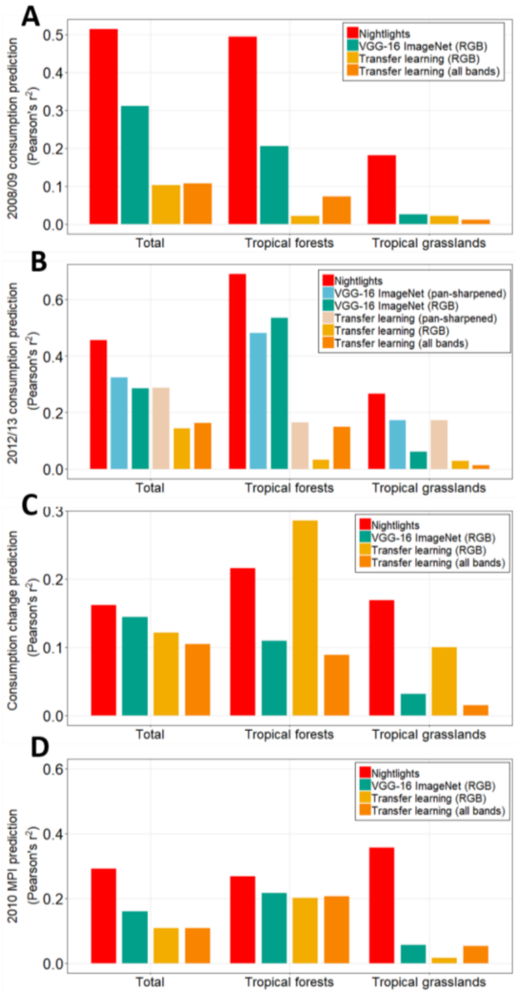

Fig. 4.2. Predictive capacity of household consumption expenditure in (A) 2008/09, (B) 2012/13, (C) household consumption change between 2008/09 and 2012/13 and (D) of multidimensional poverty in 2010 for entire Tanzania (total), and tropical forest and tropical grassland biomes, using nightlights and two different convolutional neural network (CNN) models using different imagery. We tested predictive capacity using Red-Green-Blue (RGB), 7 bands (RGB plus near-infrared, both shortwave infrared bands, and thermal infrared), and pan-sharpened imagery (only for 2013). We tested CNN models trained on ImageNet and transfer learning CNN models trained from scratch on nightlight imagery. Reported Pearson's r^2 is from five-fold cross-validation. Consumption expenditure was logged.

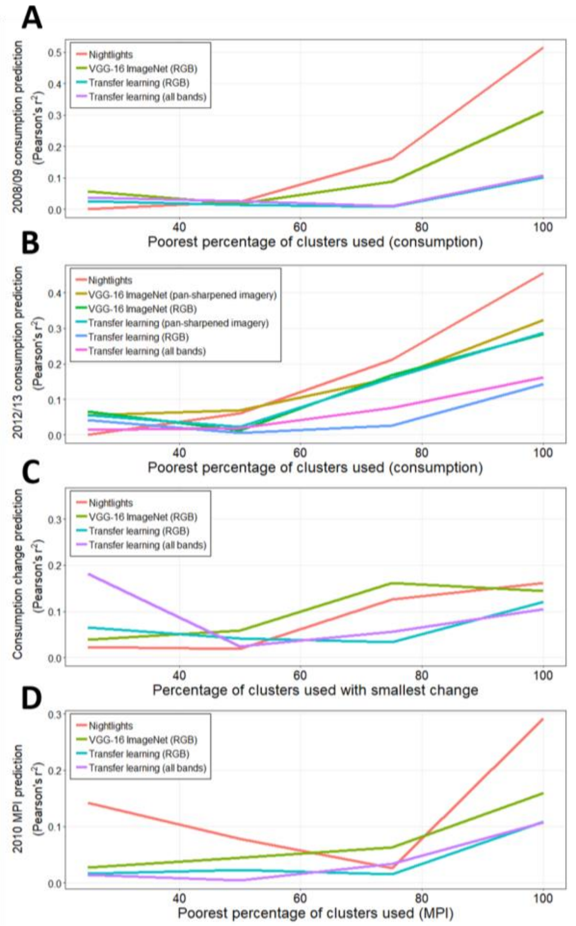

Fig. 4.3. Predictive capacity of household consumption expenditure in (A) 2008/09 and (B) 2012/13 per poorest cluster used, predictive capacity of (C) consumption change between 2008/09 and 2012/13 per percentage of clusters with smallest change, and (D) predictive capacity of the multidimensional poverty index (MPI) per poorest cluster used. Comparisons of using nightlights and two different convolutional neural network (CNN) models using different imagery. We tested predictive capacity using Red-Green-Blue (RGB), 7 bands (RGB plus near-infrared, both shortwave infrared bands, and thermal infrared), and pan-sharpened imagery (only for 2013). We tested CNN models trained on ImageNet and transfer learning CNN models trained from scratch on nightlight imagery. Reported Pearson's r^2 is from five-fold cross-validation. Consumption expenditure was logged.

Finally, we find heterogeneity depending on the CNN which we use. We find that using VGG-16 ImageNet CNN as a feature extractor generally outperforms transfer learning CNNs trained on nightlights. This could be a result of the small sample training set on which we trained the nightlight transfer learning models (only 369 images for 2008). Increasing the number of images to train the CNN will likely improve predictive capacity of transfer learning models. Yet, our transfer model CNN to assess poverty change performed better than the VGG-16 ImageNet CNN. We also find that pan-sharpened imagery leads to improved predictive power, but the difference is minimal (32% for pan-sharpened imagery and 28% for normal RGB imagery; Figure 4.1B). This could be due to overfitting issues with the pan-sharpened images that are much bigger in size. Overfitting could also be the reason that using more bands generally leads to worse predictive power. Another reason could be that CNN architectures are generally not trained on using multiple bands and could lead to biased weights, especially if some bands are not useful to predict socio-economic indicators.

To assess the robustness of these findings we compared predictive power of our models by splitting our poverty metrics into four ordinal classes. We measured the multi-class AUC and the weighted Cohen's kappa of all our models. Overall, these results confirm that CNNs can be useful to predict poverty metrics. However, many CNNs have an AUC around 0.6 which signals that they perform only slightly better than random (Figure 4.4). Nightlight as a predictor shows the highest AUC. We find that our transfer learning models sometimes slightly outperform VGG-16 based models. This suggests that the predictive power of our transfer learning models using Pearson's R^2 correlation might have been influenced by outliers. However, the differences in AUC between the models are minimal. Pan-sharpened images show the highest AUC scores. The weighted Cohen's kappa scores confirm these patterns (Figure 4.5). Our transfer learning model using 2013 pan-sharpened imagery performed best (weighted Cohen's kappa of 0.52).

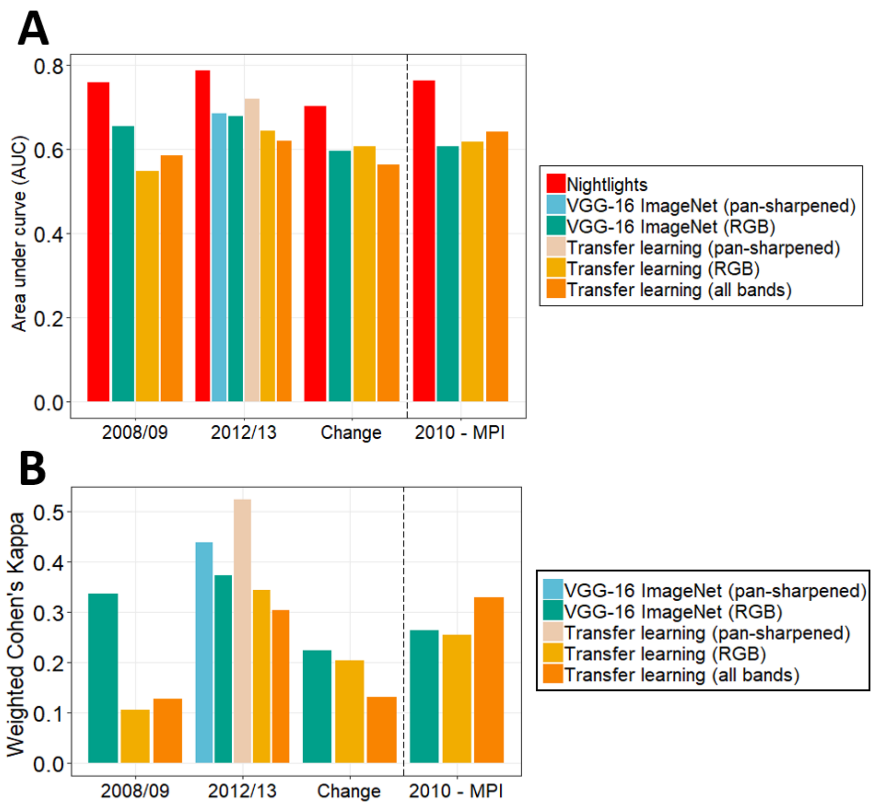

Figure 4.4. Multiclass Area Under Curve (AUC) (A) and Weighted Cohen's kappa (B) of predicting household consumption expenditure classes in 2008/09 and 2012/13, household consumption change classes from 2008/09 to 2012/13, and multidimensional poverty index (MPI) classes in 2010. Consumption, consumption change and MPI were binned in four bins. Comparisons of using nightlights and two different convolutional neural network (CNN) models using different imagery. We tested predictive capacity using Red-Green-Blue (RGB), 7 bands (RGB plus near-infrared, both shortwave infrared bands, and thermal infrared), and pan-sharpened imagery (only for 2013). We tested CNN models trained on ImageNet and transfer learning CNN models trained from scratch on nightlight imagery. Reported AUC and weighted Cohen's kappa are from five-fold cross-validation. Note that weighted Cohen's Kappa cannot be calculated for continuous variables and therefore nightlight is missing. Consumption expenditure was logged.

4.5 Discussion

Met opmerkingen [b17]: Edited discussion to add more nuance

Our study shows that CNNs trained on publicly available satellite imagery could be informative on predicting household consumption expenditure, an important indicator of poverty, in Tanzania at fine spatial scales. We achieved this accuracy despite lower resolution imagery and fuzzy survey geolocations. In contrast to published methods we used a method that can be scaled to measure poverty metrics in other countries at minimal cost and only used data sources that can be accessed freely without any licensing limitations. However, our method is not yet able to outperform previous methods, including those using freely available nightlight data. Therefore, other methods are currently preferable over our method. Imagery at lower resolution might not be able to pick up details that are important for detecting poverty, such as roof type or road quality, but accessibility to higher resolution imagery is increasing rapidly and could improve poverty predictions in the future.

In addition, we present the first attempt (to our knowledge) to measure socio-economic change over time using CNNs. Directly measuring change would be beneficial to track socio-economic indicators, but we show that current CNN methods are not yet able to predict local changes in consumption expenditure – at least using publicly available satellite imagery. However, measuring change in remote sensed imagery by using CNNs is a novel field that is still under development (Daudt et al. 2018). It is possible that our time frame was too small to detect change, but longitudinal socio-economic data collections over longer time frames are rarely implemented. One important reason for this, is that in all longitudinal surveys respondents leave the survey over time (survey attrition), gradually reducing sample sizes (Falaris 2003). For example in Tanzania, due to survey attrition a new set of households was selected from 2015 onwards in the LSMS-ISA surveys.

We show that CNN models are less capable in detecting differences in multidimensional poverty than in consumption-based poverty between clusters. Perhaps some indicators of the index, for example education, are more difficult to measure remotely than others, such as roof type (Jean et al., 2020). Future studies could assess which indicators are detectable from space and which indicators would rely on traditional survey methods.

Our study uses commonly applied CNN methods, but it is likely predictive capacity of poverty using publicly available imagery can be improved. Firstly, using more accurate geolocation source data would likely lead to better predictions (Watmough, 2019). Although CNN methods provide an opportunity to fill in data gaps, better and more data will still be required to optimize these methods. Future research could find out how much data is needed for accurate poverty predictions. Secondly, CNN methods improve quickly and newer CNN models are increasingly able to work with sparse data sources such as socio-economic data (Wang & Hebert 2016). Furthermore, CNN methods could be combined with other methods to enhance the possibility to estimate poverty. For example, our CNN models do not take into account spatial autocorrelation between images. Previous studies have shown that poverty in a location is correlated with poverty at nearby locations (Okwi et al. 2007; Graetz et al. 2018; Tusting et al. 2019). Combining CNNs with methods that incorporate spatial dependency (e.g. Gaussian processes) might further increase the capacity to estimate a variety of socio-economic indicators. For example, You et al. (2017) combined machine learning with gaussian processes to estimate crop yields, an important indicator of agricultural development.

Met opmerkingen [b18]: Example to clarify point

Our study provides three important contributions. First, it builds upon a small but growing literature of studies that apply CNN methods in remote sensing for a large range of tasks including object recognition, detection and classification (Zhu et al. 2017). Together these studies demonstrate how machine learning methods provide additional information compared to traditional remote sensing techniques. Here we use features extracted from CNNs to predict poverty, but features extracted from CNNs could provide useful information on a range of aspects (e.g. agriculture, infrastructure) that are regularly studied in impact evaluations of poverty reduction policies and programmes.

Second, our study could be replicated in other countries to test how much predictive capacity of poverty metrics varies per country. Previous studies have shown that there is considerable variation in predictive capacity of poverty metrics between countries. We show that predictive capacity can also vary significantly within a country. By testing CNN methods in multiple countries we can gain better insight into the conditions under which CNN methods perform best.

Finally, if our CNN method holds promise to be used by statistical agencies in many LMIC countries that struggle to estimate socio-economics at fine-scale resolutions and lack resources to acquire additional data. Numerous NGOs could benefit as well to improve targeting of those areas that are most in need of aid and measure the socio-economic impact of their programmes to gain insight into which programmes work and which do not (Clements et al. 2008).

4.6 References

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G.S., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, I., Harp, A., Irving, G., Isard, M., Jia, Y., Jozefowicz, R., Kaiser, L., Kudlur, M., Levenberg, J., Mane, D., Monga, R., Moore, S., Murray, D., Olah, C., Schuster, M., Shlens, J., Steiner, B., Sutskever, I., Talwar, K., Tucker, P., Vanhoucke, V., Vasudevan, V., Viegas, F., Vinyals, O., Warden, P., Wattenberg, M., Wicke, M., Yu, Y. & Zheng, X. (2016). TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems.

Alkire, S. & Foster, J. (2011). Counting and multidimensional poverty measurement. *J. Public Econ.*, 95, 476–487.

Anand, S. & Sen, A. (1997). Concepts of Human Development and Poverty: A Multidimensional Perspective. *United Nations De-velopment Program. Poverty Hum. Dev. Hum. Dev. Pap.*, 1–20.

Bengio, Y. (2009). Learning Deep Architectures for AI. *Found. Trends® Mach. Learn.*, 2, 1–127.

Blumenstock, J., Cadamuro, G. & On, R. (2015). Predicting poverty and wealth from mobile phone metadata. *Science*, 350, 1073–6.

Brockington, D., Howland, O., Loiske, V.-M., Mnzava, M. & Noe, C. (2018). Economic growth, rural assets and prosperity: exploring the implications of a 20-year record of asset growth in Tanzania. *J. Mod. Afr. Stud.*, 56, 217–243.

Chen, X. & Nordhaus, W.D. (2011). Using luminosity data as a proxy for economic statistics. *Proc. Natl. Acad. Sci. U. S. A.*, 108, 8589–8594.

Chollet, F. (2015). Keras [WWW Document]. URL https://github.com/keras-team/keras

Christie, G., Fendley, N., Wilson, J. & Mukherjee, R. (2018). Functional Map of the World.

Clements, P., Chianca, T. & Sasaki, R. (2008). Reducing world poverty by improving evaluation of development aid. *Am. J. Eval.*, 29, 195–214.

Dahl, G.E., Sainath, T.N. & Hinton, G.E. (2013). Improving deep neural networks for LVCSR using rectified linear units and dropout. In: *2013 IEEE Int. Conf. Acoust. Speech Signal Process.* IEEE, pp. 8609–8613.

Daudt, R.C., Le Saux, B., Boulch, A. & Gousseau, Y. (2018). Urban change detection for multispectral earth observation using convolutional neural networks. In: *Int. Geosci. Remote Sens. Symp.* Institute of Electrical and Electronics Engineers Inc., pp. 2115–2118.

Donahue, J., Anne Hendricks, L., Guadarrama, S., Rohrbach, M., Venugopalan, S., Saenko, K. & Darrell, T. (2015). Long-Term Recurrent Convolutional Networks for Visual Recognition and Description.

Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., Hoersch, B., Isola, C., Laberinti, P., Martimort, P., Meygret, A., Spoto, F., Sy, O., Marchese, F. & Bargellini, P. (2012). Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services. *Remote Sens. Environ.*, 120, 25–36.

Ellis, F. & Mdoe, N. (2003). Livelihoods and rural poverty reduction in Tanzania. *World Dev.*, 31, 1367–1384.

Elvidge, C.D., Baugh, K.E., Zhizhin, M. & Hsu, F.-C. (2013). Why VIIRS data are superior to DMSP for mapping nighttime lights. *Proc. Asia-Pacific Adv. Netw.*, 35, 62.

Elvidge, C.D., Sutton, P.C., Ghosh, T., Tuttle, B.T., Baugh, K.E., Bhaduri, B. & Bright, E. (2009). A global poverty map derived from satellite data. *Comput. Geosci.*, 35, 1652–1660.

Falaris, E.M. (2003). The effect of survey attrition in longitudinal surveys: Evidence from Peru, Côte d'Ivoire and Vietnam. *J. Dev. Econ.*, 70, 133–157.

Foga, S., Scaramuzza, P.L., Guo, S., Zhu, Z., Dilley, R.D., Beckmann, T., Schmidt, G.L., Dwyer, J.L., Joseph Hughes, M. & Laue, B. (2017). Cloud detection algorithm comparison and validation for operational Landsat data products. *Remote Sens. Environ.*, 194, 379–390.

Friedman, J., Hastie, T. & Tibshirani, R. (2009). glmnet: Lasso and elastic-net regularized generalized linear models. *R Packag. version*, 1.

Gilbertson, J.K., Kemp, J. & van Niekerk, A. (2017). Effect of pan-sharpening multi-temporal Landsat 8 imagery for crop type differentiation using different classification techniques. *Comput. Electron. Agric.*, 134, 151–159.

Golding, N., Burstein, R., Longbottom, J., Browne, A.J., Fullman, N., Osgood-Zimmerman, A., Earl, L., Bhatt, S., Cameron, E., Casey, D.C., Dwyer-Lindgren, L., Farag, T.H., Flaxman, A.D., Fraser, M.S., Gething, P.W., Gibson, H.S., Graetz, N., Krause, L.K., Kulikoff, X.R., Lim, S.S., Mappin, B., Morozoff, C., Reiner, R.C., Sligar, A., Smith, D.L., Wang, H., Weiss, D.J., Murray, C.J.L., Moyes, C.L. & Hay, S.I. (2017). Mapping under-5 and neonatal mortality in Africa, 2000–15: a baseline analysis for the Sustainable Development Goals. *Lancet*, 390, 2171–2182.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D. & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sens. Environ.*, 202, 18–27.

Graetz, N., Friedman, J., Osgood-Zimmerman, A., Burstein, R., Biehl, M.H., Shields, C., Mosser, J.F., Casey, D.C., Deshpande, A., Earl, L., Reiner, R.C., Ray, S.E., Fullman, N., Levine, A.J., Stubbs, R.W., Mayala, B.K., Longbottom, J., Browne, A.J., Bhatt, S., Weiss, D.J., Gething, P.W., Mokdad, A.H., Lim, S.S., Murray, C.J.L., Gakidou, E. & Hay, S.I. (2018). Mapping local variation in educational attainment across Africa. *Nature*, 555, 48–53.

Griggs, D., Stafford-Smith, M., Gaffney, O., Rockström, J., Öhman, M.C., Shyamsundar, P., Steffen, W., Glaser, G., Kanie, N. & Noble, I. (2013). Sustainable development goals for people and planet. *Nature*, 495, 305–307.

Grosh, M.E. & Glewwe, P. (1998). Data Watch: The World Bank's Living Standards Measurement Study Household Surveys Margaret E. Grosh and Paul Glewwe. *J. Econ. Perspect.*, 12, 187–196.

Hu, F., Xia, G.-S., Hu, J. & Zhang, L. (2015). Transferring Deep Convolutional Neural Networks for the Scene Classification of High-Resolution Remote Sensing Imagery. *Remote Sens.*, 7, 14680–14707.

Jean, N., Burke, M., Xie, M., Davis, W.M., Lobell, D.B. & Ermon, S. (2016). Combining satellite imagery and machine learning to predict poverty. *Science (80-.).*, 353, 790–794.

Jerven, M. (2013). *Poor numbers : how we are misled by African development statistics and what to do about it*. Cornell University Press.

Jerven, M. (2014). *Benefits and costs of the data for development targets for the Post-2015*

Development Agenda.

Ji, S., Xu, W., Yang, M. & Yu, K. (2013). 3D Convolutional Neural Networks for Human Action Recognition. *IEEE Trans. Pattern Anal. Mach. Intell.*, 35, 221–231.

Ji, S., Zhang, C., Xu, A., Shi, Y. & Duan, Y. (2018). 3D Convolutional Neural Networks for Crop Classification with Multi-Temporal Remote Sensing Images. *Remote Sens.*, 10, 75.

Karpathy, A., Toderici, G., Shetty, S., Leung, T., Sukthankar, R. & Fei-Fei, L. (2014). Large-scale Video Classification with Convolutional Neural Networks.

Kingma, D.P. & Ba, J. (2014). Adam: A Method for Stochastic Optimization.

Krizhevsky, A., Sutskever, I. & Hinton, G.E. (2012). ImageNet Classification with Deep Convolutional Neural Networks.

Kuhn, M. (2008). Building predictive models in R using the caret package. *J. Stat. Softw.*, 28, 1–26.

LeCun, Y., Bengio, Y. & Hinton, G. (2015). Deep learning. *Nature*, 521, 436–444.

Lecun, Y., Bottou, L., Bengio, Y. & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proc. IEEE*, 86, 2278–2324.

Lu, Y., Nakicenovic, N., Visbeck, M. & Stevance, A.S. (2015). Policy: Five priorities for the un Sustainable Development Goals. *Nature*.

Melaas, E.K., Friedl, M.A. & Zhu, Z. (2013). Detecting interannual variation in deciduous broadleaf forest phenology using Landsat TM/ETM+ data. *Remote Sens. Environ.*, 132, 176–185.

Nogueira, K., Penatti, O.A.B. & dos Santos, J.A. (2017). Towards better exploiting convolutional neural networks for remote sensing scene classification. *Pattern Recognit.*, 61, 539–556.

Noor, A.M., Alegana, V.A., Gething, P.W., Tatem, A.J. & Snow, R.W. (2008). Using remotely sensed night-time light as a proxy for poverty in Africa. *Popul. Health Metr.*, 6, 5.

Okwi, P.O., Ndeng'e, G., Kristjanson, P., Arunga, M., Notenbaert, A., Omolo, A., Henninger, N., Benson, T., Kariuki, P. & Owuor, J. (2007). Spatial determinants of poverty in rural Kenya. *Proc. Natl. Acad. Sci. U. S. A.*, 104, 16769–74.

Osgood-Zimmerman, A., Millear, A.I., Stubbs, R.W., Shields, C., Pickering, B. V., Earl, L., Graetz, N., Kinyoki, D.K., Ray, S.E., Bhatt, S., Browne, A.J., Burstein, R., Cameron, E., Casey, D.C., Deshpande, A., Fullman, N., Gething, P.W., Gibson, H.S., Henry, N.J., Herrero, M., Krause, L.K., Letourneau, I.D., Levine, A.J., Liu, P.Y., Longbottom, J., Mayala, B.K., Mosser, J.F., Noor, A.M., Pigott, D.M., Piwoz, E.G., Rao, P., Rawat, R., Reiner, R.C., Smith, D.L., Weiss, D.J., Wiens, K.E., Mokdad, A.H., Lim, S.S., Murray, C.J.L., Kassebaum, N.J. & Hay, S.I. (2018). Mapping child growth failure in Africa between 2000 and 2015. *Nature*, 555, 41–47.

Östberg, W., Howland, O., Mduma, J., Brockington, D., Östberg, W., Howland, O., Mduma, J. & Brockington, D. (2018). Tracing Improving Livelihoods in Rural Africa Using Local Measures of Wealth: A Case Study from Central Tanzania, 1991–2016. *Land*, 7, 44.

Pailler, S., Naidoo, R., Burgess, N.D., Freeman, O.E. & Fisher, B. (2015). Impacts of Community-Based Natural Resource Management on Wealth, Food Security and Child Health in Tanzania. *PLoS One*, 10.

Pan, S.J. & Yang, Q. (2010). A Survey on Transfer Learning. *IEEE Trans. Knowl. Data Eng.*, 22, 1345–1359.

Perez, A., Ganguli, S., Ermon, S., Azzari, G., Burke, M. & Lobell, D. (2019). Semi-Supervised Multitask Learning on Multispectral Satellite Images Using Wasserstein Generative Adversarial Networks (GANs) for Predicting Poverty.

Roy, D.P., Kovalskyy, V., Zhang, H.K., Vermote, E.F., Yan, L., Kumar, S.S. & Egorov, A. (2016). Characterization of Landsat-7 to Landsat-8 reflective wavelength and normalized difference vegetation index continuity. *Remote Sens. Environ.*, 185, 57–70.

Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Berg, A.C. & Fei-Fei, L. (2015). ImageNet Large Scale Visual Recognition Challenge. *Int. J. Comput. Vis.*, 115, 211–252.

Selomane, O., Reyers, B., Biggs, R., Tallis, H. & Polasky, S. (2015). Towards integrated social-ecological sustainability indicators: Exploring the contribution and gaps in existing global data. *Ecol. Econ.*, 118, 140–146.

Sexton, J.O., Noojipady, P., Anand, A., Song, X.P., McMahon, S., Huang, C., Feng, M., Channan, S. & Townshend, J.R. (2015). A model for the propagation of uncertainty from continuous estimates of tree cover to categorical forest cover and change. *Remote Sens. Environ.*, 156, 418–425.

Simonyan, K. & Zisserman, A. (2014). Very Deep Convolutional Networks for Large-Scale Image Recognition.

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I. & Salakhutdinov, R. (2014). Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *J. Mach. Learn. Res.*, 15, 1929–1958.

Steele, J.E., Sundsøy, P.R., Pezzulo, C., Alegana, V.A., Bird, T.J., Blumenstock, J., Bjelland, J., Engø-Monsen, K., de Montjoye, Y.-A., Iqbal, A.M., Hadiuzzaman, K.N., Lu, X., Wetter, E., Tatem, A.J. & Bengtsson, L. (2017). Mapping poverty using mobile phone and satellite data. *J. R. Soc. Interface*, 14, 20160690.

Tusting, L.S., Bisanzio, D., Alabaster, G., Cameron, E., Cibulskis, R., Davies, M., Flaxman, S., Gibson, H.S., Knudsen, J., Mbogo, C., Okumu, F.O., von Seidlein, L., Weiss, D.J., Lindsay, S.W., Gething, P.W. & Bhatt, S. (2019). Mapping changes in housing in sub-Saharan Africa from 2000 to 2015. *Nature*, 568, 391–394.

United Nations IEAG. (2014). *A World that Counts–Mobilising the Data Revolution for Sustainable Development*. *New York United Nations*.

Wang, Y.X. & Hebert, M. (2016). Learning to learn: Model regression networks for easy small sample learning. In: *Eur. Conf. Comput. Vis.* Springer Verlag, pp. 616–634.

Watmough, G.R., Atkinson, P.M., Saikia, A. & Hutton, C.W. (2016). Understanding the Evidence Base for Poverty–Environment Relationships using Remotely Sensed Satellite Data: An Example from Assam, India. *World Dev.*, 78, 188–203.

Watmough, G.R., Marcinko, C.L.J., Sullivan, C., Tschirhart, K., Mutuo, P.K., Palm, C.A. & Svenning, J.-C. (2019). Socioecologically informed use of remote sensing data to predict rural household poverty. *Proc. Natl. Acad. Sci.*, 116, 1213–1218.

Wulder, M.A., White, J.C., Woodcock, C.E., Belward, A.S., Cohen, W.B., Fosnight, E.A., Shaw, J., Masek, J.G. & Roy, D.P. (2016). The global Landsat archive: Status, consolidation, and direction. *Remote Sens. Environ.*, 185, 271–283.

Xie, M., Jean, N., Burke, M., Lobell, D. & Ermon, S. (2016). Transfer Learning from Deep

Features for Remote Sensing and Poverty Mapping. *Thirtieth AAAI Conf. Artif. Intell.*

You, J., Li, X., Low, M., Lobell, D. & Ermon, S. (2017). Deep Gaussian Process for Crop Yield Prediction Based on Remote Sensing Data. In: *AAAI'17 Proc. Thirty-First AAAI Conf. Artif. Intell.* pp. 4559–4565.

Young, N.E., Anderson, R.S., Chignell, S.M., Vorster, A.G., Lawrence, R. & Evangelista, P.H. (2017). A survival guide to Landsat preprocessing. *Ecology*, 98, 920–932.

Zhu, X.X., Tuia, D., Mou, L., Xia, G.S., Zhang, L., Xu, F. & Fraundorfer, F. (2017). Deep Learning in Remote Sensing: A Comprehensive Review and List of Resources. *IEEE Geosci. Remote Sens. Mag.*

4.7 Supplementary information

Table S4.1. The dimensions, indicators, deprivation cut-offs and weights of the multidimensional poverty index

Dimension	Indicator	Deprived if:	Weight
Health (weight 0.33)	1. Child mortality	A child under 18 years of age has died in the family in the five-year period preceding the survey	0.1667
	2. Nutrition	Any person under 70 years of age for whom there is nutritional information is undernourished	0.1667
Education (weight 0.33)	1. School attendance	Any school-aged child is not attending school up to the age at which he/she would complete class 8	0.1667
	2. Years of schooling	No household member aged 10 years or older has completed six years of schooling	0.1667
Living standards (weight 0.33)	1. Cooking fuel	The household cooks with dung, agricultural crop, shrubs, wood, charcoal or coal	0.0556
	2. Electricity	The household has no access to electricity	0.0556
	3. Source of water	The household has no access to clean water or safe drinking water is at least a 30-minute walk (roundtrip) from home	0.0556
	4. Improved Sanitation	The household does not have access to improved sanitation (flush toilet or similar)	0.0556
	5. Housing	The household has inadequate housing: the floor is of natural materials or the roof or walls are of natural or rudimentary materials	0.0556
	6. Assets	The household does not own more than one of these assets: radio, TV, telephone, computer, animal cart, bicycle, motorbike, or refrigerator, and does not own a car or truck.	0.0556

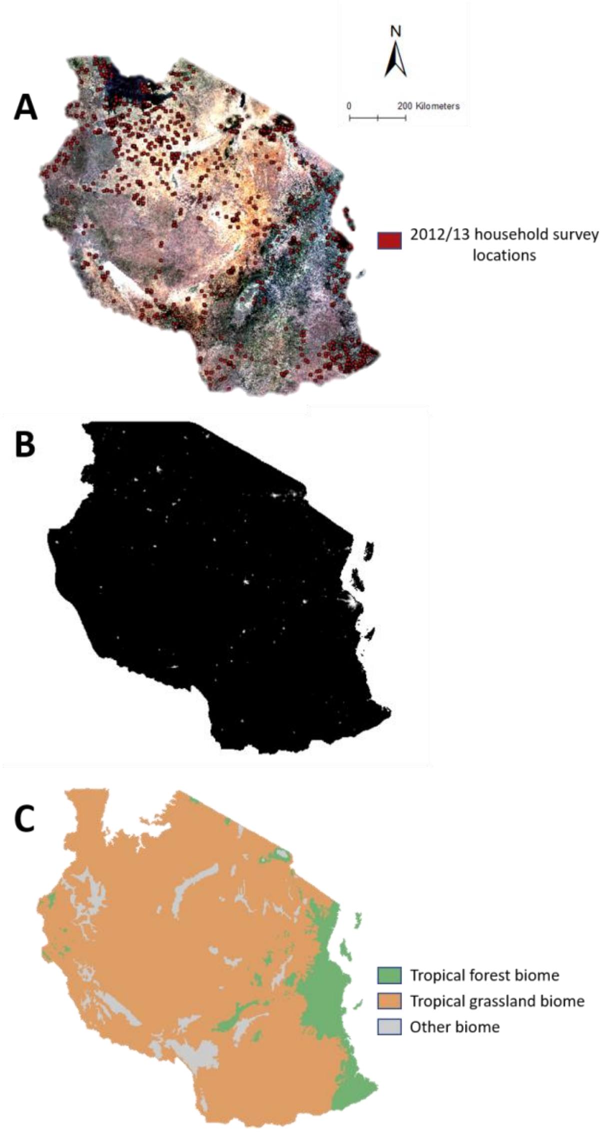

Figure S4.1 Schematic maps of (A) 2013 day-time imagery and household survey locations, (B) nightlight intensity, and (C) biomes in Tanzania. Day-time image from 2013 is an atmospherically corrected mosaic from Landsat 8 OLI/TIRS Tier 1. 2012/13 household surveys are from the Living Standards Measurement Surveys – Integrated Surveys on Agriculture (LSMS-ISA). Nightlight image is from the United States Air Force Defense Meteorological Satellite Program (DMSP).

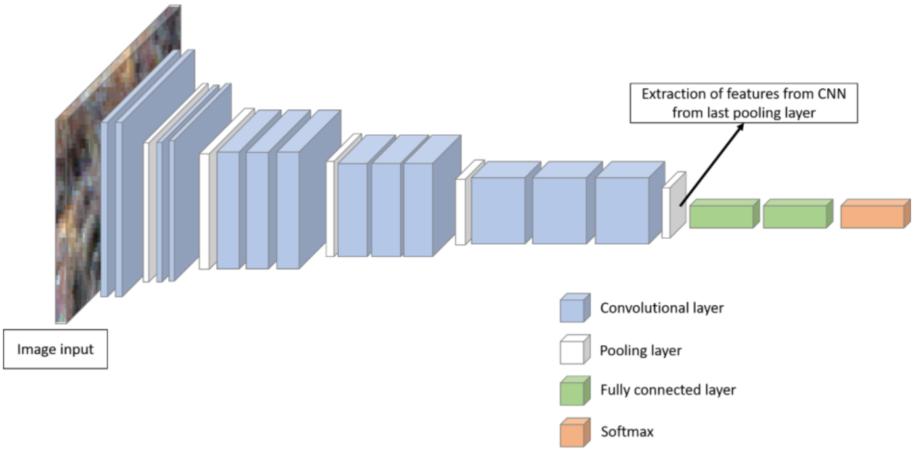

Figure S4.2. Schematic figure of the convolutional neural network (CNN) architecture, based on VGG-16. The CNN architecture consists of five blocks of convolutional layers to extract features from day-time images, followed by a pooling layer after each block to reduce dimensionality and two fully connected layers before a Softmax activation function that classifies the day-time image into a nightlight intensity category. Features can be extracted from the last pooling layer for further analyses.

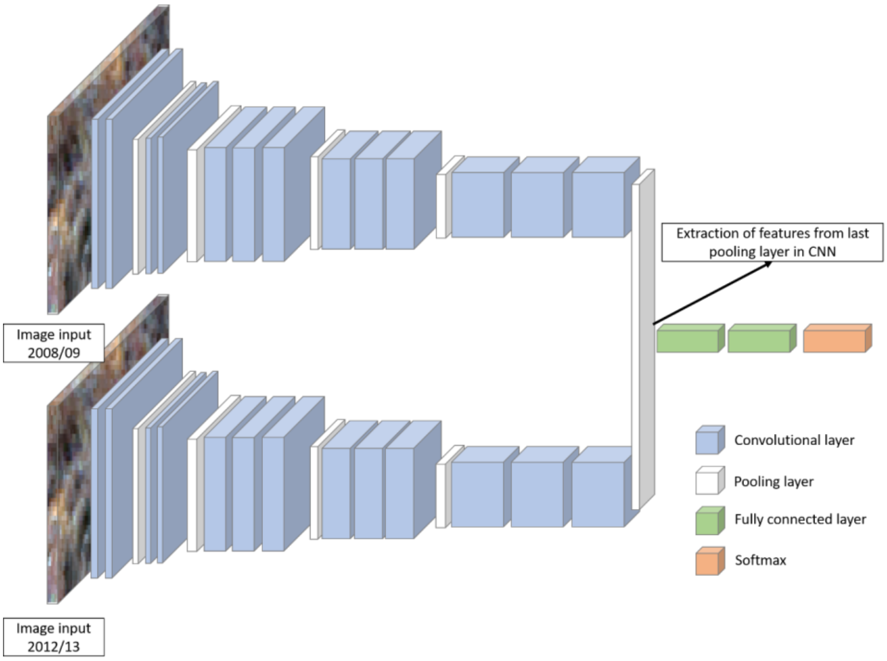

Figure S4.3. Schematic figure of the 3-dimensional convolutional neural network architecture, based on VGG-16. The 3D-CNN accepts a pair of images and extracts features from each image, then combines the features and classifies each pair of day-time images to a nightlight intensity change category. Each CNN arm consists of five blocks of convolutional layers to extract features from images, followed by a pooling layer after each block to reduce dimensionality and two fully connected layers before a Softmax activation function that classifies the image into a category. Features can be extracted from the last pooling layer for further analyses.

Chapter 5:

General discussion

5.1 Summary

Reducing poverty and halting biodiversity loss are two important sustainable development goals (Griggs et al. 2013). Currently, poverty is decreasing while biodiversity loss is still increasing (Chen & Ravallion 2010; Pimm et al. 2014). Poverty, biodiversity, and the underlying causes are linked together in coupled human and natural systems (CHANS; Liu et al. 2007). Understanding causal inference into how humans and nature interact could provide useful insights into achieving SDGs (Ferraro et al. 2018). Protected areas (PAs) are part of CHANS. PAs are established to halt biodiversity loss, but also have positive and negative impacts on people (Brockington & Wilkie 2015). Recent improvements in identifying causal inference have allowed researchers to study the environmental and socio-economic impacts of multiple PAs within a country (Andam et al. 2010; Stuart 2010). However, this approach requires large amounts of data at fine spatial grain, which are often unavailable in many of the lower and middle-income (LMIC) countries (Jean et al. 2016). Therefore, PA impact evaluations have been biased towards certain areas, and have omitted large parts of the world (e.g. sub-Saharan Africa) where poverty is most prevalent. Recently developed machine learning algorithms could potentially close this data gap and increase our understanding of PA impacts and other determinants of poverty (Jean et al. 2016).

The main aim of this thesis is to increase the understanding of environmental and socio-economic impacts of PAs. First, I explored an important mechanism, ecotourism, through which PAs impact socio-economic conditions of local people in Nepal. I found that Nepali PAs have reduced poverty and extreme poverty without exacerbating inequality. Tourism was an important factor in explaining poverty reductions, and poverty reductions were linked to tourism infrastructure (trekking routes). I then analysed impacts of SU PAs in the Brazilian Amazon where tourism impacts are expected to be much lower and where tensions between PAs and competing land uses are more tense. Furthermore, I compared sustainable use (SU) PA impacts against competing land uses (agriculture and mining) to find out which land use was best in achieving forest and socio-economic outcomes. I found that deforestation impacts of SU PAs were minimal compared to smallholder agriculture, but were much larger compared to large landholdings and mining. Socio-economic outcomes did not differ between SU PAs and agriculture, but mining sites were better at increasing income. Finally, I analysed if publicly available satellite imagery can be used to estimate poverty in Tanzania. This could be used in future studies as a basis to estimate socio-economic impacts of policies such as PAs. I found that machine learning methods (convolutional neural networks (CNNs)) were able to provide useful information to estimate consumption-based poverty in Tanzania.

In this chapter I first assess the ability of current PA impact evaluations to inform policy and management. I then show how PA impact evaluations can be improved by using examples from my thesis, and how future studies could improve PA impact evaluations. I conclude with future steps to increase the evidence base of PA impacts across multiple countries, using newly developed statistical methods.

5.2 Value and limitations of impact evaluations to inform PA policy and management

Impact evaluations have been applied in many different fields to quantitatively measure the impacts of interventions (Dehejia & Wahba 2002). Impact evaluations stem from economics but have now been applied in a variety of fields, including ecology. This thesis supports a growing literature that studies PA impacts. PA impact evaluations have already contested a few long-held beliefs about conservation. Important examples are (1) PAs can reduce poverty and are not inherently trapping people into poverty and can reduce poverty in some cases (e.g. Andam et al. 2010), (2) strict PAs are not always better in reducing biodiversity than SU PAs (Ferraro et al. 2013; Oldekop et al. 2016), and (3) even parks with few resources can reduce deforestation (Blackman et al. 2015). In general, most studies find that PAs can be moderately effective in reducing habitat loss. A much smaller evidence base suggests that PAs do not exacerbate poverty. These findings show that there is a great need to support the existing PA network, and that future PAs could have beneficial impacts for nature without exacerbating poverty although some have warned that PA expansions could also have perverse impacts (Büscher et al. 2017). This result is of great importance now discussions to set PA expansion targets post-2020 are underway (Barnes et al. 2018). Yet, there are also caveats to impact evaluations. One important limitation is that impact evaluations by definition study the past, while policy makers are concerned with the future. Whether the conditions under which the impact is observed remain true for the future is never certain. Many PA impact evaluations show that impact can be greater in areas of human pressure, but gazetting PAs in these areas is also more expensive. PA policies could reduce costs of implementing PAs by anticipating on future dynamics, for example by gazetting a PA in an area that is currently remote, but where road developments are expected.

Another important limitation of impact evaluations is caused by the definition of impact that can be confusing for PA policy makers and managers. I will explain this difficulty of definition though an example. Take two hypothetical countries A and B. Country A does not have deforestation, both inside and outside PAs. In country B deforestation is high, but inside PAs deforestation levels are much lower although deforestation still continues within PA boundaries. PA impact evaluations of countries A and B will show that PAs in country B are effective in reducing deforestation, but will find no effect of PAs in country A. Yet PAs in country A maintain more forest cover than PAs in country B. The example shows that this definition of *impact* is dependent on the context (e.g. political, climatic context) at a certain place and time. Yet PA managers generally have limited control over this context (Coad et al. 2015). Because conditions vary between countries, impact evaluation scholars have warned against simple comparisons of PA impacts between countries (Ferraro & Hanauer 2011).

Conditions not only vary between countries, but also change within a country over time. For example, many countries experience forest transitions from high deforestation to low deforestation and eventually reforestation (Meyfroidt & Lambin 2011). The measured impact of PAs against a counterfactual non-protected area is thus dependent on the time when impact is measured (Figure 5.1): a PA in a hypothetical country that successfully avoids deforestation within its borders would have limited impact when the entire country is still forested. Impact would increase over time as deforestation pressure increases, but would peak at a certain level. When previously cleared areas would reforest, then the impact of a PA would decrease and eventually become zero when reforestation increases (or even negative if you measure forest cover change). Therefore, conservation outcomes can be maximized without any impact of PAs, largely outside the influence sphere of PA managers. For example, Eklund et al. (2019) found that PAs in Madagascar with better management scores did not perform better in reducing deforestation. This shows that impact does not always equate to conservation success, but this seems contradictory for PA managers that would want to maximize the impacts of PAs. To be more useful for policy makers and PA managers impact evaluations should provide more specific insights that go beyond measuring only whether PAs work or not.

Met opmerkingen [b19]: Example to clarify point

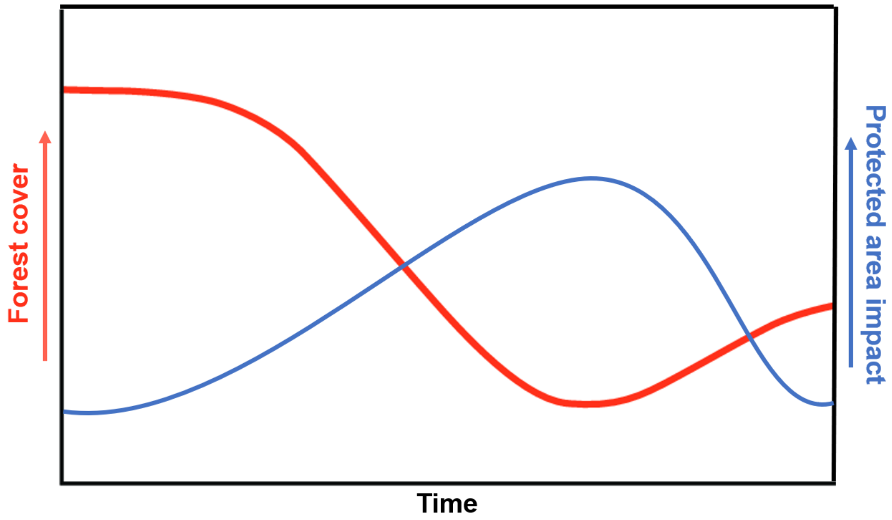

Figure 5.1. Forest cover transition and PA impact in a hypothetical country over time

5.3 Improving PA impact evaluations to inform policy and PA management

In this thesis I explored three ways in how PA impact evaluations can further improve our understanding of PA impacts and increase policy relevance. I will discuss each and explain how these can lead to a better understanding of PA impacts. I will also give recommendations on how future studies can build on these findings and improve the science on PA impacts.

First, I explored how PAs have an impact on local people. PA managers have a large set of tools to achieve positive biodiversity outcomes. Each management decision will affect biodiversity outcomes in a different way, but PA impact evaluations that only measure outcomes do not give much insight into how such outcomes were achieved. In chapter 2 I studied one important mechanism through which PAs affect social outcomes: ecotourism. I found that poverty reductions in PAs were linked to tourism and found no significant impacts on inequality. Future studies could explore the market channels through which PA policies affect ecotourism (e.g. job creation or emergence of economic opportunities; Wunder 2000). Further studies could also study how other mechanisms can affect PA impacts, for example provision of ecosystem services, hunting bans, or NGO investments.

Secondly, I studied how PAs can have heterogeneous outcomes. Understanding what causes heterogeneity in outcomes is important for policy makers to assess under which circumstances PAs are most effective, and for whom they are most effective (Ferraro et al. 2011). There are several approaches to estimating heterogeneity. One approach is to split the PAs into different subgroups. For example, in Chapter 2, I show that PAs with different tourism levels have different impacts, even though PAs with low tourism levels do not have negative poverty impacts. Other examples of subgroup analyses could split PAs into categories of funding or staffing, which are considered crucial in PA management (Waldron et al. 2017; Malcom et al. 2019), but are currently understudied in impact evaluations. Instead of subgroups, studies could assess impacts of individual PAs. Some studies have assessed individual and PA network effects on deforestation together (Eklund et al. 2019), but PAs studying socio-economic impacts have so far only focused on PA networks (largely because of sample size issues). Studying individual impacts would be useful for PA managers, but also provides insights into outliers that have exceptional positive or negative impacts (Post & Geldmann 2018), opening an avenue to study what these PAs have done well or not. Differential impacts can occur between PAs, but also within a PA (Shah & Baylis 2015). In Chapter 2, I show that poverty impacts of PAs depend on distances to trekking routes and PA entrances. Another approach can be to assess how heterogeneity is influenced by exogenous moderators, such as elevation. Ferraro et al. (2011) show that forest and poverty impacts depend on slope. At lower slopes avoided deforestation is higher, but at higher slopes poverty reduction is higher. In chapter 2 I show that poverty reductions in Nepal are not influenced by differences in elevation despite the great altitudinal profile of Nepal. Finally, heterogeneity in outcomes is important to take into consideration. I show that PAs have different effects on different poverty aspects (income, literacy, sanitation; chapter 3), and across different strata of society (poor and extremely poor, inequality; chapter 2). Many PA impact evaluations do insufficiently assess which aspects of poverty they take into account, and how these choices influence outcomes (de Lange et al. 2016).

Finally, I show how PA impact evaluations can improve by measuring PA impacts relative to other land uses and policies. There is an increased recognition that parks are not the only way to achieve positive biodiversity conservation outcomes. Community forests (Oldekop et al. 2016), indigenous territories (Garnett et al. 2018), and Other Effective Conservation Measures (OECMs; Donald et al. 2019) have been put forward as viable alternatives to protection. In some cases, these alternatives have led to better biodiversity outcomes than PAs. For example, Nolte et al. (2013) showed that indigenous territories in the Brazilian Amazon were better in reducing deforestation than protected areas. However, another alternative to protection are policies that directly affect drivers of deforestation. Such policies receive much less attention in conservation, but could also lead to significant effects on forest protection (Angelsen 2010). For example, Gabon has maintained high forest covers by influencing factors that affect agricultural productivity such as imposing heavy taxes on agricultural export, limited support to smallholders and poor road maintenance (Wunder 2005). In other countries stimulating agricultural intensification has led to reductions in deforestation by pulling away labour from extensive agriculture (Angelsen 2010). Global models that predict future biodiversity trends indicate that policies to influence agricultural impacts on reducing habitat loss can be bigger than impacts of PAs (Kok et al. 2018). To fully understand the role of PAs in achieving environmental and socio-economic outcomes impact evaluations of PAs should measure relative impacts against a wide range of policies and programmes that influence outcomes, including policies that are not area-based. Such multi-comparison evaluations could also be useful to compare socio-economic impacts. Studies have shown how PAs can have negative impacts, but these are not compared to impacts of other land uses and policies (Andam et al. 2010). It has been estimated that there are more evictions as a result of development projects than conservation interventions (Agrawal and Redford 2009). Other policies or land uses could also have negative effects on limiting access and use of ecosystem services. In Chapter 3, I show that local benefits from agriculture in the Brazilian Amazon might not be higher than benefits from SU PAs. However I also find that mining has caused positive impacts on improving income of local communities. Studies like chapter 3 can provide insight into PA impacts in relation to important drivers of biodiversity loss. It has been suggested that area-based targets for future PA expansion focus too much on the indicator instead of the overall goal (Barnes et al. 2018). Yet I would argue that only focusing on PAs is problematic in itself. A focus that is too tight on only PA risks to omit the importance of assessing drivers of biodiversity loss (Büscher et al. 2017).

Met opmerkingen [b20]: Example to clarify point

5.4 Increasing the evidence base of socio-economic impacts of PAs

The evidence base of PA impacts on deforestation has increased in recent years and now contains examples of a large number of regions across the world. In stark contrast, the evidence base of socio-economic impacts of PAs is limited to only a few countries in South-America and Asia, most of them classified as upper middle-income (UMIC; Pullin et al. 2013). There is a tremendous need for PA impact evaluations over a much wider range of countries, especially in countries with high poverty rates. This is crucial for a number of reasons. First, and most obvious, contested negative impacts of PAs on poverty should be studied in areas where poverty is most prevalent. PAs are managed in different ways in different countries, and communities in these countries have different levels of agency. Mechanisms and causal pathways that operate in upper-middle-income countries might not be feasible, or have different effects, in LMIC countries (Ferraro & Hanauer 2011). UMIC countries also have more financial resources to address conflict situations between PAs and local communities (Pullin et al. 2013). Second, more impact evaluations would allow to study PA impacts across the world. As mentioned in section 5.2 impacts are specific to the context of a certain place and time. However, if we can draw from a large number of PA impact evaluations in many different contexts, we can gain insight into how PA impacts differ across countries, across time and which factors are best in explaining PA impact.

Yet caveats remain. Despite the multi-faceted nature of poverty, impact evaluations are only able to assess certain aspects of poverty. In this thesis I assessed only non-monetary indicators of poverty in Nepal. In the Brazilian Amazon I assessed both non-monetary and monetary indicators of poverty, but I neglected important indicators such as asset ownership. Although these are important indicators of poverty, it is still possible that a separate assessment using other poverty indicators would result in different outcomes (Agrawal & Redford 2006). Just as the environmental PA impact literature is skewed towards studies on deforestation (instead of species, pollution etc.), quantitative impact evaluations on poverty are skewed towards indicators that can be more easily measured (instead of concepts that less easy to quantify, such as empowerment). Case studies of mostly single PAs have assessed impacts of interventions on indicators that are often neglected in quantitative impact evaluations. For example, a survey of four villages in North Sulawesi, close to marine protected areas, assessed the impact of a marine protected area program (Gurney et al. 2014). They focused on three poverty dimensions that are often not included in impact evaluations: security, opportunity, and empowerment. They found that the marine protected area program alleviated poverty overall, but impacts differed per individual indicator. There is evidence that such hard-to-measure indicators can be crucial in wellbeing. For example, to people living in the rural Eastern Amazon, non-monetary values (such as safety, tranquillity and community relations) were deemed more important than

Met opmerkingen [b21]: Discussion on the various definitions of poverty and what the approaches used tell us/don’t tell us, and what they may miss in other forms of poverty/wellbeing/deprivation measurement

Met opmerkingen [b22]: Example to clarify point

income (Garrett et al. 2017). Furthermore, real or perceived inequality can create conflict and impede socio-economic and biodiversity goals of PAs (West et al. 2006; Oldekop et al. 2016).

Caveats aside, because of serious financial shortfalls in LMIC countries, it is unlikely that census data - the main data source for assessing socio-economic impacts - will fill current data gaps necessary for measuring PA impacts on poverty (Jean et al. 2016). In Chapter 4, I explore whether machine learning (CNNs) can aid to fill in socio-economic data gaps using publicly available satellite imagery that can be obtained without additional costs. I find that features extracted using CNNs can provide useful information for estimating poverty at fine spatial scales. This method is easily scalable across different countries and time frames and could provide a viable approach to collecting poverty data that can be used in a wide range of applications, including PA impact evaluations. Rapid improvements in machine learning techniques and increasing accessibility of remotely sensed imagery at high resolutions will likely result in improved accuracy of poverty estimates (LeCun et al. 2015; Jean et al. 2016; Watmough et al. 2019). Other studies have used geostatistical methods to estimate a wide range of indicators of human health, education, and assets (Golding et al. 2017; Graetz et al. 2018; Osgood-Zimmerman et al. 2018; Tusting et al. 2019). The rapid increase in the availability of data on many poverty indicators could provide much better insight into the factors that affect local poverty, including PAs. Better socio-economic data from LMIC countries could improve our understanding of interactions in CHANS and create a scientific basis for achieving SDGs.

5.5 Concluding remarks

Reducing poverty and halting biodiversity loss are two important global goals that interact in multiple ways. Impact evaluations can be useful starting points to study causal inference in CHANS, including assessing environmental and socio-economic impacts of PAs. By broadening impact evaluations to also assess how and under which conditions PAs are most successful, relevance for policy and management could be increased further. Novel statistical methods such as machine learning could allow to grow the evidence base of PA impacts, especially in LMIC countries where assessments are most desired. Overall, I hope that my thesis provides useful contributions to the literature of PAs and impact evaluations.

5.6 References

Agrawal, A. & Redford, K. (2009). Conservation and displacement: An overview. *Conserv. Soc.*

Agrawal, A. & Redford, K. (2006). Poverty, development, and biodiversity conservation: Shooting in the dark? Ann Arbor MI, 48109, 647–5948.

Andam, K.S., Ferraro, P.J., Sims, K.R.E., Healy, A. & Holland, M.B. (2010). Protected areas reduced poverty in Costa Rica and Thailand. *Proc. Natl. Acad. Sci. U. S. A.*, 107, 9996–10001.

Angelsen, A. (2010). Policies for reduced deforestation and their impact on agricultural production. *Proc. Natl. Acad. Sci. U. S. A.*, 107, 19639–19644.

Barnes, M.D., Glew, L., Wyborn, C. & Craigie, I.D. (2018). Prevent perverse outcomes from global protected area policy. *Nat. Ecol. Evol.*

Blackman, A., Pfaff, A. & Robalino, J. (2015). Paper park performance: Mexico's natural protected areas in the 1990s. *Glob. Environ. Chang.*, 31, 50–61.

Brockington, D. & Wilkie, D. (2015). Protected areas and poverty. *Phil. Trans. R. Soc. B*, 370, 20140271.

Büscher, B., Fletcher, R., Brockington, D., Sandbrook, C., Adams, W.M., Campbell, L., Corson, C., Dressler, W., Duffy, R., Gray, N., Holmes, G., Kelly, A., Lunstrum, E., Ramutsindela, M. & Shanker, K. (2017). Half-Earth or Whole Earth? Radical ideas for conservation, and their implications. *ORYX*, 51, 407–410.

Chen, S. & Ravallion, M. (2010). The developing world is poorer than we thought, but no less successful in the fight against poverty. *Q. J. Econ.*, 125, 1577–1625.

Coad, L., Leverington, F., Knights, K., Geldmann, J., Eassom, A., Kapos, V., Kingston, N., de Lima, M., Zamora, C., Cuardros, I., Nolte, C., Burgess, N.D. & Hockings, M. (2015). Measuring impact of protected area management interventions: current and future use of the Global Database of Protected Area Management Effectiveness. *Philos. Trans. R. Soc. B-Biological Sci.*, 370.

Dehejia, R.H. & Wahba, S. (2002). Propensity score-matching methods for nonexperimental causal studies. *Rev. Econ. Stat.*, 84, 151–161.

Donald, P.F., Buchanan, G.M., Balmford, A., Bingham, H., Couturier, A.R., de la Rosa, G.E., Gacheru, P., Herzog, S.K., Jathar, G., Kingston, N., Marnewick, D., Maurer, G., Reaney, L., Shmygaleva, T., Sklyarenko, S., Stevens, C.M.D. & Butchart, S.H.M. (2019). The prevalence, characteristics and effectiveness of Aichi Target 11′s "other effective area-based conservation measures" (OECMs) in Key Biodiversity Areas. *Conserv. Lett.*

Eklund, J., Coad, L., Geldmann, J. & Cabeza, M. (2019). What constitutes a useful measure of protected area effectiveness? A case study of management inputs and protected area impacts in Madagascar. *Conserv. Sci. Pract.*

Ferraro, P.J. & Hanauer, M.M. (2011). Protecting Ecosystems and Alleviating Poverty with Parks and Reserves: "Win-Win" or Tradeoffs? *Environ. Resour. Econ.*, 48, 269–286.

Ferraro, P.J. & Hanauer, M.M. (2015). Through what mechanisms do protected areas affect environmental and social outcomes? *Philos. Trans. R. Soc. B-Biological Sci.*, 370.

Ferraro, P.J., Hanauer, M.M., Miteva, D.A., Javier Canavire-Bacarreza, G., Pattanayak, S.K. & Sims, K.R.E. (2013). More strictly protected areas are not necessarily more protective: evidence from Bolivia, Costa Rica, Indonesia, and Thailand. *Environ. Res. Lett.*, 8.

Ferraro, P.J., Hanauer, M.M. & Sims, K.R.E. (2011). Conditions associated with protected area success in conservation and poverty reduction. *Proc. Natl. Acad. Sci. U. S. A.*, 108, 13913–13918.

Ferraro, P.J., Sanchirico, J.N. & Smith, M.D. (2018). Causal inference in coupled human and natural systems. *Proc. Natl. Acad. Sci. U. S. A.*, 201805563.

Garnett, S.T., Burgess, N.D., Fa, J.E., Fernández-Llamazares, Á., Molnár, Z., Robinson, C.J., Watson, J.E.M., Zander, K.K., Austin, B., Brondizio, E.S., Collier, N.F., Duncan, T., Ellis, E., Geyle, H., Jackson, M. V., Jonas, H., Malmer, P., McGowan, B., Sivongxay, A. & Leiper, I. (2018). A spatial overview of the global importance of Indigenous lands for conservation. *Nat. Sustain.*, 1, 369–374.

Garrett, R.D., Gardner, T.A., Morello, T.F., Marchand, S., Barlow, J., Ezzine de Blas, D., Ferreira, J., Lees, A.C. & Parry, L. (2017). Explaining the persistence of low income and environmentally degrading land uses in the Brazilian Amazon. Ecol. Soc., 22, art27.

Golding, N., Burstein, R., Longbottom, J., Browne, A.J., Fullman, N., Osgood-Zimmerman, A., Earl, L., Bhatt, S., Cameron, E., Casey, D.C., Dwyer-Lindgren, L., Farag, T.H., Flaxman, A.D., Fraser, M.S., Gething, P.W., Gibson, H.S., Graetz, N., Krause, L.K., Kulikoff, X.R., Lim, S.S., Mappin, B., Morozoff, C., Reiner, R.C., Sligar, A., Smith, D.L., Wang, H., Weiss, D.J., Murray, C.J.L., Moyes, C.L. & Hay, S.I. (2017). Mapping under-5 and neonatal mortality in Africa, 2000–15: a baseline analysis for the Sustainable Development Goals. *Lancet*, 390, 2171–2182.

Graetz, N., Friedman, J., Osgood-Zimmerman, A., Burstein, R., Biehl, M.H., Shields, C., Mosser, J.F., Casey, D.C., Deshpande, A., Earl, L., Reiner, R.C., Ray, S.E., Fullman, N., Levine, A.J., Stubbs, R.W., Mayala, B.K., Longbottom, J., Browne, A.J., Bhatt, S., Weiss, D.J., Gething, P.W., Mokdad, A.H., Lim, S.S., Murray, C.J.L., Gakidou, E. & Hay, S.I. (2018). Mapping local variation in educational attainment across Africa. *Nature*, 555, 48–53.

Griggs, D., Stafford-Smith, M., Gaffney, O., Rockström, J., Öhman, M.C., Shyamsundar, P., Steffen, W., Glaser, G., Kanie, N. & Noble, I. (2013). Sustainable development goals for people and planet. *Nature*, 495, 305–307.

Gurney, G.G., Cinner, J., Ban, N.C., Pressey, R.L., Pollnac, R., Campbell, S.J., Tasidjawa, S. & Setiawan, F. (2014). Poverty and protected areas: An evaluation of a marine integrated conservation and development project in Indonesia. Glob. Environ. Chang., 26, 98–107.

Jean, N., Burke, M., Xie, M., Davis, W.M., Lobell, D.B. & Ermon, S. (2016). Combining satellite imagery and machine learning to predict poverty. *Science (80-.).*, 353, 790–794.

Kok, M.T.J., Alkemade, R., Bakkenes, M., van Eerdt, M., Janse, J., Mandryk, M., Kram, T., Lazarova, T., Meijer, J., van Oorschot, M., Westhoek, H., van der Zagt, R., van der Berg, M., van der Esch, S., Prins, A.G. & van Vuuren, D.P. (2018). Pathways for agriculture and forestry to contribute to terrestrial biodiversity conservation: A global scenario-study. *Biol. Conserv.*, 221, 137–150.

de Lange, E., Woodhouse, E. & Milner-Gulland, E.J. (2016). Approaches Used to Evaluate the Social Impacts of Protected Areas. *Conserv. Lett.*, 9, 327–333.

LeCun, Y., Bengio, Y. & Hinton, G. (2015). Deep learning. *Nature*, 521, 436–444.

Liu, J., Dietz, T., Carpenter, S.R., Folke, C., Alberti, M., Redman, C.L., Schneider, S.H., Ostrom, E., Pell, A.N., Lubchenco, J., Taylor, W.W., Ouyang, Z., Deadman, P., Kratz, T. & Provencher, W. (2007). Coupled human and natural systems. *Ambio*.

Malcom, J., Schwartz, M.W., Evansen, M., Ripple, W.J., Polasky, S., Gerber, L.R., Lovejoy, T.E., Talbot, L.M. & Miller, J.R.B. (2019). Solve the biodiversity crisis with funding. *Science (80-.).*, 365, 1256.1-1256.

Meyfroidt, P. & Lambin, E.F. (2011). Global Forest Transition: Prospects for an End to Deforestation. *Annu. Rev. Environ. Resour.*, 36, 343–371.

Nolte, C., Agrawal, A., Silvius, K.M. & Soares-Filho, B.S. (2013). Governance regime and location influence avoided deforestation success of protected areas in the Brazilian Amazon. *Proc. Natl. Acad. Sci. U. S. A.*, 110, 4956–4961.

Oldekop, J.A., Holmes, G., Harris, W.E. & Evans, K.L. (2016). A global assessment of the social and conservation outcomes of protected areas. *Conserv. Biol.*, 30, 133–141.

Osgood-Zimmerman, A., Millear, A.I., Stubbs, R.W., Shields, C., Pickering, B. V., Earl, L., Graetz, N., Kinyoki, D.K., Ray, S.E., Bhatt, S., Browne, A.J., Burstein, R., Cameron, E., Casey, D.C., Deshpande, A., Fullman, N., Gething, P.W., Gibson, H.S., Henry, N.J., Herrero, M., Krause, L.K., Letourneau, I.D., Levine, A.J., Liu, P.Y., Longbottom, J., Mayala, B.K., Mosser, J.F., Noor, A.M., Pigott, D.M., Piwoz, E.G., Rao, P., Rawat, R., Reiner, R.C., Smith, D.L., Weiss, D.J., Wiens, K.E., Mokdad, A.H., Lim, S.S., Murray, C.J.L., Kassebaum, N.J. & Hay, S.I. (2018). Mapping child growth failure in Africa between 2000 and 2015. *Nature*, 555, 41–47.

Pimm, S.L., Jenkins, C.N., Abell, R., Brooks, T.M., Gittleman, J.L., Joppa, L.N., Raven, P.H., Roberts, C.M. & Sexton, J.O. (2014). The biodiversity of species and their rates of extinction, distribution, and protection. *Science (80-.).*

Post, G. & Geldmann, J. (2018). Exceptional responders in conservation. *Conserv. Biol.*, 32, 576–583.

Pullin, A.S., Bangpan, M., Dalrymple, S., Dickson, K., Haddaway, N.R., Healey, J.R., Hauari, H., Hockley, N., Jones, J.P.G. & Knight, T. (2013). Human well-being impacts of terrestrial protected areas. *Environ. Evid.*, 2, 19.

Shah, P. & Baylis, K. (2015). Evaluating heterogeneous conservation effects of forest protection in Indonesia. *PLoS One*, 10.

Stuart, E.A. (2010). Matching Methods for Causal Inference: A Review and a Look Forward. *Stat. Sci.*, 25, 1–21.

Tusting, L.S., Bisanzio, D., Alabaster, G., Cameron, E., Cibulskis, R., Davies, M., Flaxman, S., Gibson, H.S., Knudsen, J., Mbogo, C., Okumu, F.O., von Seidlein, L., Weiss, D.J., Lindsay, S.W., Gething, P.W. & Bhatt, S. (2019). Mapping changes in housing in sub-Saharan Africa from 2000 to 2015. *Nature*, 568, 391–394.

Waldron, A., Miller, D.C., Redding, D., Mooers, A., Kuhn, T.S., Nibbelink, N., Roberts, J.T., Tobias, J.A. & Gittleman, J.L. (2017). Reductions in global biodiversity loss predicted from conservation spending. *Nature*, 551, 364–367.

Watmough, G.R., Marcinko, C.L.J., Sullivan, C., Tschirhart, K., Mutuo, P.K., Palm, C.A. & Svenning, J.-C. (2019). Socioecologically informed use of remote sensing data to predict rural household poverty. *Proc. Natl. Acad. Sci.*, 116, 1213–1218.

West, P., Igoe, J. & Brockington, D. (2006). Parks and peoples: The social impact of protected areas. Annu. Rev. Anthropol., 35, 251–277.

Wunder, S. (2000). Ecotourism and economic incentives - an empirical approach. *Ecol. Econ.*, 32, 465–479.

Wunder, S. (2005). Macroeconomic change, competitiveness and timber production: A five-country comparison. *World Dev.*, 33, 65–86.