GOING GEOSPATIAL WITH IMPACT EVALUATIONS

WEBINAR PRESENTATION TRANSCRIPT

JULY 15, 2015

This document was produced for review by the United States Agency for International Development. It was prepared by the Feed the Future Knowledge-Driven Agricultural Development (KDAD) project. The views expressed are those of the author and do not represent the views of the United States Agency for International Development or the United States Government.

Contents

Presenters

Dr. Ariel BenYishay, AidData

Dr. Mark Buntaine, University of California Santa Barbara

Dr. Daniel Runfola, AidData

Tania Alfonso, USAID (Moderator)

Presentation

Tania Alfonso: Ok. So hello again everybody. We're going to start. I would say good morning to about – I forgot the poll is not there anymore but about maybe 60 percent of you are tuning in from North America so good morning to you. And I'm very impressed and very grateful for those of you who are joining us from other continents. I realize it's later. For some of you it's late at night so welcome everyone. I first heard about geospatial impact evaluations probably a couple of years ago. It's really exciting to learn all the details and potential of this technology.

So we're going to see three examples of geospatial impact evaluations this morning which themselves are interesting for people involved in land tenure, environmental conservation and malaria sectors. But you should also look out for basic insights around impact evaluations, sort of what they are, what it means to identify treatment and comparison areas in a rigorous way and to be able to compare them to make conclusions around impact. So basically impact evaluations 101. And then learn about when and how it is possible to use it using geospatial information.

So we'll have presentations from three presenters and then we'll have time for Q&A. What I would ask is that as the Learning Lab sponsors already said please type any questions you have throughout the presentation into the chat box on the lower right and then there are some people available that can answer them on the spot if it's something that can be answered with a quick answer. And if it requires sort of a more drawn out answer then we will leave that for the end for our Q&A section of this presentation. So there will be plenty of time to ask questions and for people to respond to those questions.

So our three presenters are Dr. Ariel BenYishay who is a chief economist at AidData and he's an assistant professor of economics at the College of William and Mary and Dr. Daniel Runfola is also at the College of William and Mary and he is a geospatial scientist. And Dr. Mark Buntaine is the assistant professor at the University of California Santa Barbara. And I have heard all of them present before. It's very exciting. So I will leave it to the presenters to get started.

Ariel BenYishay: Great. Thanks Tania and thank you everyone for joining us. This is Ariel BenYishay. We really appreciate the chance to tell you more about what we are calling geospatial impact evaluation. As Tania noted this is an idea that's been germinating for quite a while and has some application in a few fields and now really kind of taking hold

and getting applied across the board to a wider variety of sectors. So we think about geospatial impact evaluation at AidData and as a broader membership of AidData affiliated researchers as essentially the use of spatial information on program activities.

And that spatial information on program activities can be things like the boundaries of particular parcels of land that are demarcated or where tenure rights are improved for citizens in a particular country. That could be the catchment area for a health clinic distributing anti-malaria bed nets or it could be the wider jurisdiction for a municipality in say Latin America where governance improvements are being made. So the reason that geospatial impact evaluation is really kind of coming into its own is essentially because we have much, much better resolution, much higher resolution data on these program activities spatial features. So again we now understand to a much higher degree of precision exactly where these activities are taking place and where those affects that we might estimate are located.

That is really crucial because we can then merge that with a variety of other georeferenced data sets after the outcomes of these programs. Many of these other data sets are also increasingly higher resolution. Some of them are the kind of typical survey data that we often think about in the evaluation world, household surveys capturing things like childhood diseases and mortality that are now increasingly being georeferenced. That is where the coordinates or general location at which each survey is connected is now known and provided to evaluators and other users. So that is one set of outcomes. Those surveys again capture outcomes on health but they also increasingly capture outcomes on other – in other domains such as governance and democracy or trust outcomes, social mobilization, some of those features as well.

There's another class of data sets that is again available in increasingly higher resolutions and that is a set that is remotely sensed via satellites and other sensors and that, those data sets that led us to more domains like conservation and deforestation, they also allow us to think about economic activity outcomes more broadly such as through the use of nighttime lights data that again reflect the current economic activities to a relatively high degree of precision spatially. And the reason that these, putting these two types of data sets together, that high resolution data on program activities and the high resolution outcomes merging them together is that – the reason that is so powerful is that it really allows us to think about causal attribution seriously typically through the use of quasi experimental designs like matching estimators or through the use of fixed effects with essentially controls for the differences across locations that are

consistent over time or through the use of other quasi experimental techniques like regression discontinuity as well.

So we are going to show you essentially three examples. Two of those focusing on land rights and their impact on deforestation. And the third example thinks critically about how to do this in the health sector. As I mentioned before there is a growing array of examples in this case. So a growing number of these types of evaluations in the governance sector and the post conflict domains as well as in education and other domains as well. So we won't have time to talk about all of these today but if these are sectors that you're working on or are particularly interested in we're happy to talk further with you specifically about them.

So with that I'm going to turn it over to our AidData research consortium member, Mark Buntaine who is going to walk you through a really exciting piece of work that he and several other research consortium members completed on Ecuador. So Mark, I'm going to turn it over to you.

Mark Buntaine: Thank you Ariel. So this is Mark Buntaine. Again I'm located at the University of California at Santa Barbara. And today I wanted to present to you all collaborative work with Stuart Hamilton and Marco Millones that we completed together with the USAID mission in Ecuador. So this is joint work that has been done with that mission over the last two years. This work was recently published open access in the *Journal of Global Environmental Change* and if you would like to see the research more in detail you can grab a copy of the paper at our group's website GEOIE.org and see some more details if we don't have time to get to them today.

Ok. So I wanted to tell you a bit of a story about how our particular research group got into the world of geospatial impact evaluation. We were doing a review of evaluations regarding forest land tenure. So there are donors around the world who are engaged in making sure that people have secure access to forested areas. Those programs are attempting to make forest degradation less likely by giving people ownership over the longer term management of resources. But what we had typically found is an evaluation looked something like this. You sent out a staff member to the field to conduct interviews for a period of one or two weeks and then an evaluation for upwards of in this case a $400 million project was written on the basis of some more anecdotal information that came from interviews.

Now that type of data is useful but we thought well what we really want to see is the outcomes of these types of interventions and we leverage new types of remotely sensed data to do that. Now before I

get into the specific case I want to talk just for a moment about the suite of programs or interventions where this type of approach might be useful. Our group is particularly interested in evaluating interventions that are associated with changes in land cover. And we don't think this is limited to the forestry sector that you see in the upper left. But any time that you can see as the result of a program whether it's agricultural productivity, whether it's the development of infrastructure, whether it's the development of aquaculture. You can see changes to landscapes. We can then use these types of techniques to evaluate impacts for programs that have information on where they're located.

So let me tell you a little bit about our particular program in Ecuador. We are working in Morona-Santiago, Ecuador which is in the southeastern part of the country. Over a period of six years the USAID Ecuador mission engaged in a program to offer indigenous communities secure tenure to particular plots of land. And so on the left hand panel here you see in the black checked areas the intervention sites. There were 53 communities where the mission went out and found and facilitated the secure tenure of these lands. Now in yellow you see the deforestation pattern over the period of the program. And so when you look at this you might see one of two things. The first thing you might see is that well, it looks like the program areas have occurred in places where there's less deforestation. That could mean that the program was largely effective at decreasing deforestation over the project time period.

The other possibility is the program areas are located where it's simply less likely for deforestation to occur because they're further away from roads. They are less populated. And so what we want to do then is find comparison areas so that we get a good sense about whether or not the program was actually having the impact on deforestation or not. In other words we want to take the program area and compare apples to apples. If we think that distance to roads is an important aspect of land cover change what we need to do is we need to find other areas that are also far away from roads in order to make this type of attribution that Ariel talked about in his introduction.

Let me give you a very simple example about how this works in practice. So let's say for a moment that you want to take a treated plot, there's some programming that's happening there and you want to compare it to a non-treated plot and look at the difference between those two places. So in this case the treated plot is in green. Well this treated plot is close to a road. It is on a high slope and it is far away from a town. So if we search through the landscape we can look for comparison plots and some of them will match and some of them will not. So this plot that I just brought up is close to a town so

it is not a good match. This plot right here is on a very flat area. So it might be deforested more easily because it's productive for agriculture. This plot is further away from a road that I just brought up. What we can do is we can search across the landscape to this fourth plot on the right hand side of the landscape which it the same in terms of the observable characteristics. It's on a high slope. It's close to a road and it's far away from a town. Now there are many more variables than this but this illustrates the basic search that we do across the landscape using advanced algorithms to find the correct matched sets.

So in essence that's a brief introduction to the process that's happening here. In our particular case we have a lot of different variables that we measure, all of them being georeferenced that drive deforestation. And so what we want to see is are there control or comparison areas that are similar in observable characteristics to the program areas? Are they – do they have similar trajectories? Are they the same distance away from infrastructure? Do they have the same indigenous land classification and protected area status? Are they the same elevation and slope, etcetera?

So when we do that this is what we obtain. So this is the first part of the study where we look at simple legalization. The first part of the program offered only legalization without additional land management training. And so the program areas on this map are displayed in red. When we use this algorithm to find comparison areas across the province what we end up with are the areas in yellow. These areas are similar on all of the observable characteristics to the areas that occurred for the program. Now what do we find when we compare these areas in terms of how much deforestation occurred in the five years after the program? So on the vertical axis –

Mark Buntaine: On the vertical – on the vertical axis here we have a, the amount of deforestation that has occurred over five years. If as in bar A we do not do this matching so we do not find the good comparison areas, it looks like the program was very effective. And this makes intuitive sense because the program if we remember the map happened in those areas where less deforestation occurred. After we do the matching however it does not seem that there has been an impact of this program on deforestation during the first five years of the program. Now we wanted to examine a second part of the program. This had legalization interventions plus community training for community management of natural resources. Again we take the program areas in red and we match them to comparison plots which are displayed in yellow. We find the same pattern which is if we just did this kind of analysis without finding good comparison areas it appears that there is less deforestation over five years in the program areas as compared to the non-program areas. However after we look

about the – after we look at the matched comparison sets we do not find a difference in deforestation during the program time period.

Now future directions, this was important for the Ecuador mission because they were at a point where they were deciding on the broad classes of interventions that they wanted to undertake in the forestry sector. And land tenure interventions were one among many possible programs. And so this kind of information provided evidence that in the near term it does not appear that land tenure interventions cause a decrease in deforestation. Perhaps if the goal is to have a very short term impact other types of programming areas would be preferred. So we think that this has broad applicability across different project areas. Anywhere where we know the location both of the programs and we have data that can be sensed from satellites that give us an indication of effectiveness.

The other thing that I should say is this looked at land tenure interventions over five years. And many people have said to us I think fairly that maybe effects only happen after 10 or 20 years in these areas. But typical evaluation and monitoring budgets do not have 10 or 20 years to follow up on programming. And so we think that there's great potential to match these types of evaluation tools to programs where the theory of change is longer term. And then finally we think that there is an opportunity to combine this type of geospatial impact evaluation with more traditional evaluation methods to understand mechanisms.

So this type of 30,000 foot view of evaluation can tell you if a program had a particular impact. It has less to say about why it had that impact. And we think that there's particular potential to work with implementing agencies like USAID to combine these tools that give you broader impact assessments together with more traditional evaluation. So thanks before I turn it back over to Ariel. We could not have done this work without a collaboration of the USAID Ecuador mission and all of the implementing partners of that mission on the ground so we're very grateful for that. And again if you wanted to see the paper you can grab it as GEOIE.org. Thanks very much and I'll turn it over.

Daniel Runfola: The general idea here 'cause this is a joint project by the World Bank and KFW, the German development bank – in 1988 the Brazilian government as part of its constitution promised to give legal land rights to a number of different communities within the Amazon. As a part of this $25 million investment that was made by both the World Bank and the German development bank they set a number of different objectives including giving just local governance rights and civil rights. But in addition they were interested to see if giving legal

ownership over these lands helps to influence the rates of deforestation.

To set the stage for this a little bit more broadly – so one thing that we did know about the area going in is that right now land tenure security was not widely shown to reduce deforestation. Right now indigenous control irrespective of legal ownership was already shown to relate to these lower deforestation rates. So really the question that we're trying to answer here is once legal ownership is conferred to these different individuals if when that ownership is conferred you end up seeing either lower rates in deforestation or if you see an increased rate of reforestation in some of these areas. So this project was enacted between 1995 and 2008. 181 community lands were considered for demarcation. Of those 106 community lands were actually demarcated. This in terms of spatial scope was an enormous project. So we ended up – or it ended up covering 38 million hectares, almost 35 percent of all indigenous lands in the Amazon.

To give you an idea of how we're defining these terms demarcation, this is actually the recognition by the ministry of justice that the lands were actually the borders of these lands. So there were a number of other different stages in the demarcation process so regularization is one of those. That is the entry into the municipal, state and federal registries. When this period of demarcation actually occurred varied by community between 1995 and 2008 with the median year of demarcation being 2001. An additional feature of the project was some of these communities had further support for boundary enforcement. So to give you an idea the data that we were using to examine this question of if community demarcation actually had an impact on deforestation. Working with KFW we were able to receive the boundaries of the community lands which I'll show you in a moment. So the actual geospatial area that each community covered. Additionally we had administrative data on the demarcation dates so when the demarcation occurred for each one of those boundaries.

With that data we were then able to merge in information from the NASA land long term data record. This is a very long term data series so going back all the way to 1982 we have daily observations that give us information on what the forest cover in four kilometer to five kilometer areas across the entire Amazon look like. This is using something called the normalized difference vegetation index which I'm happy to talk about in more detail later. And that ranges from zero to one. So a zero indicates very rocky or barren terrain whereas a one indicates dense forest. Ultimately we collapsed those daily measurements into annual measurements of the maximum NDVI observed and the mean NDVI observed for every single community for that entire time range so from 1982 to 2010.

In conjunction with that outcome measure we had a number of different covariates that we used. These included climate covariates so precipitation and temperature which we had for a very long time period from 1900 all the way to 2010. We have topology information so elevation and slope from the shuttle radar topography mission, distance to rivers across the Amazon and then interpolated population estimates across the entire Amazon. So this is within the environment sector all of the variants that we thought might influence our outcomes measure. But I know a lot of you are from other sectors and so just to give you a broad window into some of the types of data that we're able to get from these broad global sources we also have information on nighttime lights so you can use nighttime lights to measure kind of economic growth or infrastructure growth. We have a lot of information on conflict from ACLED, SCAD, and the early conflict warning system. And then we also have a lot of information which you'll hear more about in a moment on health outcomes from the DHS and Afrobarometer.

But for this particular project looking at forest cover this was the set of covariates that we used. The empirical methodology we used was propensity score matching. So Mark just showed you a good example of how this works in practice where you are trying to identify the treatment and control communities that are very, very similar to one another along all observable covariates. We demarcated two different treatment and control groups. One of these was early treatment versus late treatment. So if they were treated in the first half of the program, 1995 to 2001 versus the second half, 2001 to 2008. And then the second treatment and control group we constructed was if they were ever demarcated or never demarcated.

The method we used then to model the causal effects was a fixed effect model. So that allowed us to control for the time-invariant community un-observables in this particular case. This is the map of the communities across the Amazon. The lighter colors indicate areas that received earlier demarcation whereas the darker colors indicate areas that received later demarcation. There was a demarcation strategy that was in place at the beginning of the project. How closely that was followed is unclear at this point but you can see there's a general preferential early treatment of some very, very large communities in the area. This is a graph illustrating the NDVI trends for every single one of our communities.

So as the, on the y axis it goes from zero to one. One indicates very heavy levels of forestation so a lot of forest in the area. And then lower values would indicate less forest in the area. As you can see – so within these indigenous communities from 1982 to 2014 there is a general upward trend of forestation. So we receive higher levels of

NDVI over time in these communities. This is of course counter to the Amazon writ large where you see trends of deforestation but runs alongside what we observed in the literature. That is that indigenous communities are in fact good at protecting their forests.

This is the same chart but looking at the late treatment versus early treatment. So the red line on this chart is communities that received an earlier treatment from 1995 to 2001 and the blue line is those that received an earlier treatment so that was from 1995 to 2001. Apologies. The red line is 2001 to 2008. And again you can see in either case we observe that upward trend. So these are after the matching, the actual findings that we had. So the very, very first bar you see at the top there is the treatment effect. In this case we're looking at a cross sectional model examining the maximum NDVI between every community between 1995 and 2010. So similar to Mark's study we do not see strong indications of an effect of the treatment.

Similarly, so this is the early versus late model and this is looking at the outcome between 1995 and 2001 so the question that we're examining here is in the short term, so in the first half of the study if we observed an outcome based on the treatments. So in the earlier time stage we do not see clear evidence of the outcome influenced by the treatment. And then this is the cross sectional results for the later period so between 2001 and 2010. And again we see non-significant impacts associated with the treatment effect.

Finally what you see here – we were actually because we have the temporal information on when the treatment occurred. We had yearly information on the outcome of NDVI and we had yearly information of both temperature and precipitation over time we were able to compile all of this information along with our community fixed effects to look at a panel model of if we believe two different types of treatment have – or if we have evidence that two different types of treatments had an impact. The first of these is the date of demarcation so in the panel model the year that demarcation occurred received a one and all years after that received a one. And then we did the same thing for enforcement. So communities, some communities received additional funding to actual enforce their demarcated boundaries. Again in the panel case you see in both instances we did not see strong evidence of an impact of these treatments.

So preliminarily what we can conclude right now is that we did not find clear or robust evidence of differences in deforestation attributable to the PPTAL project. There are much lower rates of deforestation on indigenous lands in the cross-sections but these may not be related to land tenure status of these lands. I think that in

general what we observed and as you saw on the NDVI trend graphs is that these indigenous lands tend to have an upswing in the amount of forestation over time irrespective of the treatments that we tested in this study.

There are a lot of next steps in future research here. The first of these is that we really need to identify high pressure communities so this is communities where there was a priority, a higher chance of deforestation occurring where treatment effects may be larger. Additionally we would like to disaggregate to smaller units of analysis adding more precision from finder covariates. So for example right now we examined this study at the community level but we have NDVI level all the way down to about the five kilometer pixel scale. Additionally we have administrative data on the criteria for the timing of the community demarcation. So there are some selection concerns still within the model but we have the information to adjust for that.

Finally we would like to expand the control groups by including communities that never entered PPTAL. So there are a number of communities that were simply never demarcated but we do not have all of the information on those at this time. By identifying these communities we will be able to expand our group and try to identify in our never versus ever demarcated model with more precision. I'm going to go ahead and pass it off to Ariel who is going to be talking to you a little bit about a similar project evaluating malaria aid.

Ariel BenYishay: So to those of you who may not work in the land tenure sector or that don't often interact with forest cover outcomes- well, now I want to focus on an application of geospatial impact evaluation that looks at human outcomes in terms of childhood survival and looks at the health sector specifically. So this work that I'll be presenting is the most preliminary of the various projects that we're presenting today. It is joint work together with several researchers at other universities also active in the health and malaria sector. The basic premise of this project is essentially an evaluation of a decade long investment by the World Bank in the health sector in the Democratic Republic of Congo.

Essentially between 2005 and last year the World Bank invested upwards of $200 million in this sector in the DRC. Much of it spent on fighting malaria. Now the World Bank is not the only actor in the DRC investing heavily in malaria. The Global Fund and other actors are there as well but this represents a very large portion of all spending on malaria in the last decade in the DRC. That spending funded a large number of interventions including treatment for pregnant women, first line meds for malaria and importantly scaling up coverage of long life insecticide treated bed nets via a government

led mass distribution campaign. That mass distribution campaign spread the use of bed nets across the country and it is partly that that we will be essentially evaluating.

Using data that Aid Data recently provided we are able to map the locations of intervention through this project in the DRC to a very high level of precision. In fact this particular malaria project is one that we were able to map to the highest degree of precision possible in our data. So the number of locations that you see displayed here are mapped to very high precision. In the map that I'm showing here World Bank locations are displayed in red. Now the World Bank investments were made over several specific projects and so the lighter dots are representing locations funded through the first project. Darker dots in red display locations funded through the second wave of the project. In blue we've mapped clusters of the demographic and health survey. This is as many of you know a reputable and standardized data collection effort that is carried out across the world and importantly also provides the geographic coordinates of the survey cluster locations.

We are then able to map the cluster locations from the two most recent rounds of the demographic and health surveys carried out in the DRC and show how they overlay with these sites which World Bank interventions were funded. As you can see the coverage of the DHS data which are displayed in blue dots here on this map is reasonably wide across the country over these two waves and has reasonable overlap with the sites at which the World Bank interventions took place as well. The basic evaluation question for this effort is to be able to address the child mortality impact of these investments. Our question essentially asks whether child mortality decreased after a project is activated and whether that mortality decrease differentially for children located closer to our project site relative to a reference or comparison group of children located relatively farther from that project site.

By making these comparisons both across time and over distances from these project sites we're able to control for both time varying effects and differential outcomes both closer and farther away from these project sites. We're essentially using a survival analysis model that adopts the differences-in-differences approach, essentially again controlling for district-level fixed effects and for trends at the district level. For those of you who like regression models and slides I've included a basic model here. We are essentially looking at the differential child survival across children who are born closer and farther away from a 2005 project site after that 2005 project site is active. And similarly for the 2011 wave sorry project funding we're looking at children who are born closer and farther away from those

sites both before and after 2011, that gives us both that comparison across distance and over time.

Differences-in-differences models of these types may still be plagued with selection bias if for example the World Bank reasonably could predict where child survival was likely to increase or sorry or decrease as a result of malaria trends. And thus invest differentially in places that were likely to see changes over time. Thus we're attempting to use one additional source of quasi-experimental variation. That is one based on time varying malaria ecology. The basic premise here is that malaria risks due to the natural environment vary over time and over space in ways that are not easily predictable by policy-makers or World Bank project planners. And thus we can use those as another source of essentially natural or quasi-experimental variation to assess where childhood survival is affected by the malaria ecology as well.

One of the reasons this approach may reasonably work in this setting is that the malaria ecology data is also available at a relatively fine or granular level. So we have as I'm showing on this slide the malaria ecology both over time and this is one snapshot of the long term malaria ecology across the DRC but also on a monthly and annual basis. So essentially we can compare years in which the weather was reasonably dry and reasonably cool so that malaria bearing mosquitoes did not breed at such high rates and the risks for malaria in those years were reasonably small and use that as yet another source of comparison. And then compare those to years in which again the weather may have been hotter and wetter so such malaria breeding, sorry, mosquito breeding was much higher and transmission of malaria across the population was likely to be much higher and use those particular windows to then examine the effectiveness of the interventions by using that difference-in-difference model I showed before.

One other value of using this malaria ecology index is that in the malaria sector specifically one of the big kind of remaining questions for the sector is what share of the community needs to treated for a treatment effect to materialize. So whether interventions need to be focused on raising community treatment above a certain threshold in order for a treatment effect to materialize or are there thresholds or disproportionate gains from treatment across the community. One of the reasons that this particular index lets us do that is that it essentially varies the load in the population from malaria and allows us to assess treatment effects across varying modes so that we can essentially assimilate what portion of the population is being treated and then assess whether the treatment effects vary over that varying portion of the population.

So with that we're going to turn to just a few practical questions. We've talked about several cases of this geospatial impact evaluation applications but we'd like to talk to you also about what broadly you might think about as you might prepare for doing one of these in the particular context that you're facing. First of all just a few advantages that you probably have picked up on as we've been talking. But one aspect of geospatial impact evaluation is that we can often recover baseline data using existing georeferenced data. The demographic and health surveys again that I've just shown include child and child birth and death outcomes for all adult women in a household that are interviewed so that we can trace out child survival even the child survival that's taken place reasonably far back in time.

As both Dan and Mark showed we can use remotely sensed data that are available often before a particular intervention took place as well to recover these baselines. So that is one big advantage as we see it to geospatial impact evaluation its application in cases where a given project may not have had baseline data collected. Another advantage as we see it is that it can be accomplished reasonably quickly and often less expensively than a large impact evaluation that demands customized data collection and demands treatments to be administered prospectively, that is in the future, rather than retrospectively as in these cases.

These advantages do have some needs. The need is really the spatial variation. So as we've just shown you these projects require reasonably precise information on where project activities took place and that could be either in the form of community boundaries or parcel boundaries or specifically the sites at which again bed nets are distributed and applications to the governance sector. That can be locations and municipal offices where particular interventions are carried out and in the education sector that can be the locations of schools where upgrades or investments are carried out. That is essentially one of the big needs. For those of us, for those of you looking to engage with Aid Data to help kind of think through the particular challenges and opportunities in your context I should note that Kristina Kempkey and Brian Bingham are our representatives and they are excellent starting points for the conversation.

As you contact Kristina and Brian you can start thinking about kind of the first set of questions that we are likely to have which is essentially again how precisely are the locations or sites of intervention known. So as I mentioned the spatial variation need is one crucial component to geospatial impact evaluation. And so understanding the both the actual location but also what data is available about the locations or sites of the project is the first crucial question. We should also know something about how the locations or sites were chosen, about how that rollout across these sites was

planned so that we can control for whether this was reasonably quasi experimentally – sorry whether a quasi-experimental design is reasonable or not.

And then finally the outcomes of interest so as many of you have indicated in your questions in the chat box which we'll address in just a moment the particular outcomes of interest tie very closely to particular types of data that are required and those data and their particular levels of precision and timeframes that they're available over will really shape whether a geospatial evaluation is feasible for the particular outcomes that you're addressing.

So those are the very first set of questions that we can really start talking about as we talk with each of you about your particular questions or opportunities. At this point we'll turn it back over to Tania and she can help direct us through the many excellent questions that you all have posed and we will do our best to answer them.

[End of Audio]

Q&A

Tania Alfonso: Thank you so much. This is really exciting. I think this has beaten the record for number of attendees to a Learning Lab webinar. This is really exciting. A couple sort of – I wanted to put in a plug for USAID Learning Lab which was what hosted this webinar. You should continue the conversation through the discussion forum that I just mentioned on that chat box because I'm not sure that we're going to be able to answer every single question in the next half hour. And the Evaluation Interest Group on Learning Lab is just a great resource for if you'd like to continue conversations about evaluation in general and about this particular type of impact evaluation. So I'll probably mention it one more time before we conclude the call.

I think what I want to do is to maybe go sort of in reverse order from the presenters. So the last thing that Ariel talked about was just general like how do we use GIS, how do we access AidData? And so there's some questions about GIS, the use of GIS in general and I'm going to look at those first so the questions are sort of in the left-hand side of the chat box. And I think Ariel and Dan should answer these. And then we'll go back to the specific evaluations that were discussed in these presentations. So I guess Ariel and Dan maybe if you answer the first three questions.

Ariel BenYishay: Sure. So on this topic of how GIS can be used on governance project one example that we have of a project just launched in Columbia is a municipal governance improvement project where the government with the help of several additional actors is attempting to improve local governance. And in those cases we can merge data that is being collected through surveys of the population in these areas and use that to link to these particular sites at which interventions are carried out. Again the spatial aspects or the GIS aspects allow us to consider the distance from the intervention sites as well as kind of the boundaries of the intervention sites. So it allows us to construct comparison groups in some of these cases based on the individuals that are located relatively far from where a given intervention is carried out or individuals who are just beyond the jurisdiction of that intervention.

So in the cases where an intervention involves say travelling to a municipal center to interact with the government again a comparison group might be a set of the population that's located relatively farther away from it, from that municipal center. Or in cases where the municipal government again is acting across the entire municipality we can use spatial data on the locations of survey respondents so then construct a comparison group of individuals located just beyond the jurisdiction or in other jurisdictions not simultaneously treated.

So that is one way that we can use the locational information or the GIS aspect for governance projects as well. I should mention that we are actively working now on geocoding the Afrobarometer data sets for the first six rounds, some of which have been geocoded in the past but not to a common standard. And AidData itself is attempting to geocode these to a common standard together with Afrobarometer so that we can have a standardized data set available at least at this point for much of Africa covering many of these governance domain questions. Dan, I'll let you talk about this, the remote sensing question.

Daniel Runfola: Sure. So –

Tania Alfonso: Dan if you could read the questions out loud I think that would be helpful.

Daniel Runfola: Absolutely. One of the questions was how can evaluators get access to some of the great remote sensing data that is being generated? This is a huge challenge. So the data sets that we were using we were talking about daily images for 30 years so you're up in the terabytes of data which is really unusable for traditional audiences. One of the things that we're hard at work at here is actually generating a fairly simple to use online interface that will allow anyone to extract that data to administrative or other units of interest so being able to upload a shape file or select an administrative unit that you are studying and then automatically download those time series. We expect that will be rolled out in kind of early forms by the end of the summer approaching into the early fall and that should be the best answer to that question. Right now unfortunately just the scope of the data makes it very, very difficult for kind of common users to access.

The other one, so I'll briefly answer number three as well. Yeah. so would it be possible for the speakers to share simple perhaps open source tools that can allow someone to do basic comparisons such as this and that was in reference to the propensity square matching and approaches that we saw earlier. One thing that we are also actively working on and we already have out in the wild on GitHub for those of you that do any kind of development that will soon be on the CRAN and broader repositories we have in all the programs, so R is a statistical program that allows you to do some geospatial analysis.

That program is open source and allows you to do exactly these types of analysis. Now it's not simple yet and I won't claim it is yet. This is a relatively young methodology and so the tools to do this are still fairly nascent. But at the same time it allows anyone to start exploring some of these tools. I think that one thing that we would

be particularly interested in is learning more about what types of tools would make your jobs easier as this methodology begins to evolve. And with that I think I'll pass it off. It looks like number four is a question for Mark.

Mark Buntaine: Ok. Sorry for that. We've got my audio back on. Thank you. So the question is what happens to the treated areas that could not be matched in the Ecuador study and perhaps more broadly what happens if you can't find good matches to do this type of evaluation? So there are actually lots of different approaches to matching comparison groups. So the work that Dan presented uses a model to find those comparison sites that are equally likely to be treated. That's one approach and it requires more of a one for one match which might make this problem more significant. The approach we use is actually a comparison at the level of the overall observations so we are not matching one treatment plot to one control plot. We are looking at distribution averages across all of the variables we care about within the complete set of observations.

So I suppose the short answer is there are different ways to do matching that enable this to be possible. Now it could be the case that you just can't find good match sets and then you have a couple of different options. One is you use further model adjustments to make sure that the control observations are similar to the treated observations. The second option is that you look entirely at those treated plots that can be matched well and that offer some basis for comparison. Both have their drawbacks. But the overall message that I want to put forward is in doing this type of impact evaluation and particularly matching there are a lot of different types of techniques that need to be tailored to the application that they are applied to. And with that I'm turning it back over to Dan I believe.

Daniel Runfola: Great. So to read the next question a lot of the evaluation depends heavily on the measure of deforestation. What approach was taken to measure rates of deforestation? And then there were a number of very technical questions along these lines as well so I'll try to kind of knock a lot of these out in one answer. What we ended up using was the NASA land long term data record, LTDR which is a process combination of AADHRR and MODIS data that goes back to around circa 1982. The spatial resolution of that, they standardize it to about a 0.05 decimal degree so that's roughly five to five-and-a-half kilometers throughout the Amazon. That is NDVI but it is heavily processed NDVI. So they account for things like cloud cover.

And the other big advantage is because we have daily measurements we are able to overcome a lot of problems that are really, really big in the Amazon. So clouds are very, very common in the Amazon and

so it's hard to see from a satellite where trees are. If you have daily pictures however you have a much better chance of actually capturing the variation in forest cover on the ground. We have also been testing a few other products that are helpful for measuring this so the GIMS product we looked at and that extends backwards in time to a similar time horizon.

One big question is why not look at Landsat or why not go back farther. So the fundamental answer to that question is the processing costs become very, very high. So we have Landsat imagery. In theory we could go back all the way to the 1970s but right now the costs of doing that would be phenomenal so even the MODIS data which is 250 meter resolution raw, then we're talking about trillions and trillions of pixels of data. If you go to Landsat you're talking about even more than that. So while there are a number of initiatives currently ongoing to classify the Landsat data, it is not, there is not currently a publicly available dataset that has good coverage over the Amazon.

The other big challenge in going back farther is that because Landsat only takes one picture every 16 days it's frequently plagued by cloud cover issues which AADHRR or MODIS, though they are coarser spatial resolutions they do have fine temporal resolutions. With that I think I will turn it over. I believe the next one is for Mark asking could you explain more about fragmentation in and around treatment areas.

Mark Buntaine: Yeah. And maybe I will just follow up a little bit on Dan's response to the measures of deforestation. So in our case we take, we use a data set put together by Hanson and Colleagues that measures vegetation at five meters height. So there are more and more data products that make it possible to do this type of analysis around the world using common metrics. That's one thing I want to explain. The other thing is that it is now possible to validate those types of broadly available open sourse metrics with some of the higher resolution data that Dan measures. So in our case we use MODIS data and measure of enhanced vegetation index averaged over annual basis to validate some of the results that come out of the Hanson data and we have all of that written up in very lengthy technical appendixes if anyone wants to get down into the weeds.

The next question was can you explain more about fragmentation in and around the treatment areas? So in our particular study the average rate of deforestation is about two to three percent per year around this area. We did not look specifically at fragmentation as an outcome but I would like to respond to one point which is we know that there are frontier effects in these kinds of forest changes. That is if a plot is deforested the plots that are near to it are more likely to be

deforested in the future. And so what we do is a general methodology is not only to match on all of those observable variables that I put a list up of but also to say around the plot what is the level of deforestation?

So let's find comparison plots that near to the plot have the same levels of deforestation. And furthermore what is the trajectory of deforestation coming up to the treatment? So if a program goes into effect in 2002 what we do is we look back over the years before 2002 and we try to find areas that have the same trajectory of nearby fragmentation to those areas that fall under the program. So in terms of the more general application of this methodology we think that there's a real need to think about these types of issues particularly to make sure that you're understanding spillover effects and that you're understanding some of these other spatial dynamics that drive the results.

And then maybe I'll just speak briefly to point seven before turning it over to Ariel. So this is now that their research team has gathered data on deforestation could research be used to gather other types of outcomes like income or access to credit. This is absolutely the direction that our group would like to head. And we think that there are real gains to integrating this type of evaluation work into the traditional evaluation work that is already happening. Often times at low cost. So if there are opportunities to collaborate to build this type of data set where you're looking not only at forest cover change or other types of land use change but also to complement that with real household level data we think that that is potentially powerful. It just requires a commitment to work with an implementing partner that already has access to those types of data. Ariel?

Ariel BenYishay: Yeah. I think most of that Mark is exactly spot on. So I think in some cases where survey data is also available on some of these populations from alternative sources that can be a reasonably good proxy. And in certain types of environments again we can use other remote sensing, remotely sensed outcome measures as well. Again lights or other measures of economic activities to merge in as well. And it partly depends on the level of I guess capacity and existing data in the country as well of course.

So where census data is available to a reasonably fine level or were other national household surveys are also available to relatively fine levels those outcomes are certainly kind of within the range of the doable. Dan, let me ask you to answer question number eight and then we have a number of other questions that we'll try to address too in the time that we have remaining.

Daniel Runfola: Sure so I will briefly engage number eight. When do geospatial evaluations – sorry. How do you reduce the likelihood of ecological fallacy when linking different data sets based on geography? So briefly just to define the concerns of ecological fallacy this is when you have very coarse resolution data and you are going to find the resolution areas and it's closely related to the modifiable area unit problem which I know I saw a few other questions about which is where you have units of observation that may not be measured using the same spatial units of analysis. So for example different administrative units or census units as contrasted to government units.

In both of those cases we treat those are sources of spatial uncertainty. So you can see actually on the slide that we're looking at right now the first set of questions we're exploring is how precisely our locations and sites known. That applies to all of the different data sets that we are working with. And so if you have coarse resolution information it's actually a great source of spatial uncertainty. Fortunately there are a lot of people that have worked hard to allow us to model that type of information.

This is actually closely related to another question that was submitted asking if there were any additional steps particular to spatial impact evaluations required to protect personally identifiable information. In those contexts you can actually take that data and you can make it spatially less precise and then you can incorporate information on the uncertainty around that data in your modeling efforts. So there's a lot of different ways to engage with this challenge and we are actively working on solving those issues. With that I will pass it back to Ariel on question nine.

Ariel BenYishay: So are any HIV/AIDS studies ongoing that are using GIS to inform those studies? I should say that there aren't many. I do not know of any that are ongoing at the moment that are essentially attempting to use GIS as a source of quasi-experimental variation or to be able to do causal attribution in quite the way that we're describing them. But that is certainly not because of lack of applicability. In some sense we are in active discussions with a number of groups who have reasonably precise information on activities carried out to combat HIV/AIDS and attempting to have a marriage, for example the DHS data with other data sets such as vital statistics or national clinic or facility surveys that essentially provide some of these outcome measures again at pretty fine levels of precision.

The spatial aspects really also let us bring in other data as well, roads information for example in cases where HIV/AIDS epidemics are believed to spread along transportation routes or are believed to be correlated with other baseline or background characteristics like

poverty or governance. Those are instances in which we can exactly look at those cases. So there is a rich kind of applications in the HIV/AIDS world that we're really starting to explore. Dan mentioned briefly in response to the question about protecting personally identifiable information.

And just as you would in cases where the spatial aspects are not known one needs to essentially insure that there is sufficient coarseness in all of the characteristics that one cannot link any particular set of responses or set of information to any one user or any one respondent that is essentially the application in the GIS world. You have to think about a richer set of variables that could be used to identify any one observation and then assess in both the data and if it is a sample then in the real world what is the total population identifiable by this same data set that we have in our possession.

I should say that the DHS has been releasing the georeferenced survey data for quite a while now and they have their own standardized ways to do this. Part of it is by essentially masking the locations introducing small degrees of error into the locations so that you cannot easily re-identify the cluster locations. So there are kind of established ways to do this and those ways are also being constantly improved. Mark, I'll let you answer number eleven.

Mark Buntaine: Thanks Ariel. So question eleven is is it possible to crowd source or make an intuitive interface for project recipients to self-report in order to do evaluation over longer time periods? So I think this is a really interesting suggestion in an area that's ripe for collaboration for different organizations that are doing this type of work. My sense and I know that AidData and our group was just at a workshop put on by USAID not too long ago where we talked about long term evaluation is that a lot of theories of change happen over a number of years and that makes evaluation particularly challenging because it's just not the case that there's a continuing field presence for years and years after projects are happening.

But let's say projects that are in the agricultural sector. Do transitions of crops brought about through extension services have lasting impacts on income and do they spill over to other nearby farms? These are really important questions that get to the core of these programs but are very difficult to monitor over longer periods of time. So I think that there's great opportunity to build the kinds of reporting systems that are being suggested here together with more of the macro land cover change data that one could see have the cropping system shift from single season to two crops in a season types of systems as a result of programs for let's say climate change adaptation. So this is an area that's ripe for collaboration and we think a real knowledge need as well.

Daniel Runfola: Great. So I'll answer just briefly question 12. Just a follow up question on the NDVI. What additional aspects were taken into account to determine if a project, the project was the main cause of the changes on the community land and not the other aspects, say climates changes? Was the climatic data and other data integrated to account for this? So absolutely. We had around 15 different covariates that we looked at. The climatic data in particular we actually have globally for every single month from 1900 to 2010 both temperature and precipitation measures. And we have that at a thoroughly fine spatial resolution so we were able to then integrate that with our community level data sets. With that I will pass it off to number 13 for Ariel.

Ariel BenYishay: Are the presenters familiar with any examples of layering GIS with land value control trial impact evaluations focused on household level labor or income outcomes? So I can talk about one example that I've been personally involved in and that is an impact evaluation in Malawi of an agricultural extension program where the intervention was meant to have farmers adopt a more sustainable and climate resilient technology. And in that case we were able to merge using GIS data merge climate conditions that varied across Malawi in our intervention years and some of the locations rainfall was particularly poor in our intervention years and in some not. And we were able to use that to really test the mechanism underlying the particular intervention again that this technology that the farmers were adopting was really climate resilient by merging in GIS data on climate.

So that is kind of one way to use that as a cross cutting measure of again natural or a quasi-experimental variation even in an RCT case where the intervention of interest is randomly varied across the population. It looks like there's quite a few other questions. We'll take just a moment to see if we can try to figure out a reasonable way to address them. For those of you – I should apologize in advance for those of you whose questions we did not fully address. Please do email Christina and we will do our best to address those via email as well.

Tania Alfonso: Right. And again I want to put in a plug for Learning Lab and also I think the interest to join Learning Lab as well and we can continue the conversation there. There's an Evaluation Interest Group that we have multiple different discussions this would only be one of them. And it's something that we can share opinions, experiences, challenges, opportunities. I've also seen a great amount of exchanges, emails and contact information within the chat box. It seems like there are many people are kind of connecting with each

other and this also sets a record for the webinar. It's probably the most interactive so this is super, super exciting.

Mark Buntaine: So this is Mark Buntaine. I'll take a shot at this longer question which is given that two or three of the projects evaluated were related to land tenure could we talk a little bit about the study design? So I'm just bringing up a separate slide here to that I hope will talk about some of the limitations. So the first point in this question is are there limits to this study design mainly have we controlled for all of the other drivers for deforestation besides the...impact evaluation. And the honest answer is we do the best we can and we think we're there but we always have to be careful to make sure that we're not over claiming the comparison group. So here is the approach that we have taken. In the left of this slide we talk about difference-in-differences.

This looks at let's say that the treated groups in a solid line have a greater increase in forest cover as compared to control groups. Well the challenge of just a simple difference-in-differences approach to analysis is maybe we've missed something else that's not the same between the treated and the control. So the assumption that we make that drives a lot of this analysis is if we get control sites that are the same in terms of the outcome at the point of the program start and we follow them back in time to show that they're on the same trajectory up to the program start we're confident that those unobserved factors that we haven't accounted for, if they were truly important they would show up in this trajectory and pretreatment outcome. So this is the assumption that we make to deal with this problem.

To the second question is there a limitation to defining treatment as the presence of the program alone rather than perceptions of land tenure we can think of this as for those of you who are in the impact evaluation world as an intent to treat kind of analysis. That we can't show which communities took up the intervention best but since the goal of the implementer is simply to deploy the intervention what we would like to do is say "Given that this is how it was deployed on average do you have an effect?"

In terms of number three do the findings lead project designers to consider the importance of incorporating new economic incentives? My answer would be I hope so. This is often a challenge for evaluators which is we come up with results and we would like those results to be used and that's why we seek partnership with organizations like USAID to make sure that the evaluations that are happening are useful and are designed in ways that can answer ongoing questions. So thanks for the question.

Ariel BenYishay: Thanks for the excellent answer Mark. We don't really have much to add from our side at this moment. I should just say one more time that if you do have additional questions or if you want to have a conversation about the particular evaluation opportunities that you are facing or challenges that you have we're happy to do that and the best way to get in touch with all of us is probably through Kristina Kempkey and Brian Bingham who's email addresses are provided. And again I'd encourage you to get engaged in the Learning Lab more broadly. Tania, is there anything else that you want to add at this point? I think that we had a chance to cover most questions. If there are still major questions please do add them to the chat box.

Tania Alfonso: Yeah. I would request again that join the conversation on Learning Lab. Join the Evaluation Interest Group. There's no reason why these discussions have to end right now. Again the chat has been amazing. There's been amazing participation and I hate to cut anybody off. So again there is now a link being put on the chat box for continuing the discussion so please go ahead and continue if you can. We're going to move over to some poll questions in a moment so please take the time to answer the poll questions. Those of you who are left.

And again I would like to thank everyone who joined. I'd like to thank the presenters for their long involvement for sitting and answering all the questions patiently and I want to thank the 260 odd people who joined, some of you in very unforgiving time zones. So I'm really gratified that the level of interest was really incredible and I'm really happy about that. So again I just want to thank everyone who joined in, who asked questions, who answered questions. This has been great. So please, please check the polls. Please go ahead and answer them really quickly. I'm going to answer them myself as well. And then after that I will look forward to continuing the conversation.

[End of Audio]