Projeto # 15
Effects of Road Investments on forest in Brazil's Amazon

Alexander Pfaff, Earth Institute, Columbia University, USA
Juan Robalino Earth Institute, Columbia University, USA
Eustáquio Reis, IPEA-RJ
Cláudio Bohrer, UFF

Overview

Tropical forests receive immense attention as locations where economic development and ecosystem conservation regularly clash. The few major areas of tropical forest are prized for species habitat and carbon storage. However, clearing is ongoing, and it may continue given the local importance of development and difficulty of regulation. In this setting, road networks may be the only 'lines in the sand' to shape the spatial pattern of forest.

The Amazon Basin has exceptionally high biodiversity, its forests may play roles in global carbon and hydrological cycles, and its regional deforestation is longstanding. In the early 1960s, the government began to build roads linking the Amazon to other parts of Brazil. Since then, over 16% of Brazil's pre-Columbian Amazonian forest has disappeared and ongoing forest loss is currently around 2.4 million hectares per year.

In 2000, a Brazilian-government initiative originally known as 'Avança Brasil' (Advance Brazil) called for paving an additional 7500 km of highways in Amazonia. Some plans are widely debated, such as paving the 900 km BR-163 highway, which would facilitate soy exports.[1] How much forest would be lost from such investments?

Not surprisingly, empirical research that addresses this question has been done. Pfaff (1999), e.g., based on econometric analysis of deforestation in the 1970s and 1980s, provides a numerical association between greater road density within a county and higher deforestation in that county as well as higher deforestation in the neighboring counties.[2] Per such 'conventional wisdom', Laurance et al. 2001 presented scenarios of forest loss in the Amazon based on assumptions about the impacts of infrastructure on forest cover.[3] These suggested that, fully implemented, Avança Brasil would lead to the loss of 28% of the pre-Columbian forest by 2020 if "optimistic" and a loss of 42% if "non-optimistic".

This contrasts sharply with Andersen et al. 2002's evaluation of census data for the Amazon.[4] Concerning the impacts of roads on forest clearing, they give dramatically different conclusions. They found that the impact of roads depends critically on the level of previous forest clearing and, most dramatically, that with substantial prior clearing new roads *reduce* the rate of deforestation. This, they suggest, is because roads in these circumstances focus local development and draw it away from remaining forest tracts.

The proposed research program will re-evaluate these questions using far more precise data. Andersen et al. employed census data for Brazilian counties (municípios), resulting in a sample of 257 county units (quite large on average although varied in size) for the entire Brazilian Amazon. We will employ census tracts to provide 20+ times as many observations. This and data from the 1970s forward will permit empirical gains.

Analyses of Impacts on Overall Deforestation

Single Period, Impacts of Road Investments in Road Locations

Analyses can be done for each period of time observed (in terms of forest data from satellite remote sensing the periods will be 1976 - 1987, 1986 - 1992 and 1992 - 2000). For example, we

[1] D. C. Nepstad *et al., Science* **295**, 629 (2002).
[2] Many find similar results for other places, such as Chomitz and Gray 1996 and Pfaff and Sanchez 2004.
[3] This assessment stimulated considerable debate, often focused upon the assumptions concerning impacts of new roads on forest cover. For debate, see e.g., do Amaral 2001, Goidanich 2001, Silveira 2001 and Weber. For alternative derivations of parameters from observed deforestation see, e.g., Stavins and Jaffe 1990 and Nelson and Hellerstein 1995 plus, for the Amazon, Reis and Guzman 1992 and Pfaff 1999.
[4] Note that their book addresses a wide range of issues, including the economic benefits from forest conversion. It is not focused on roads. Their work advanced methodologically the study of Amazon deforestation. For instance, their use of lagged road changes to explain deforestation is applied here. Further, they put significant effort into applying an objective method for choosing among regressions.

have results for the impacts of 1968 – 1975 road investments on the 1976 – 1987 deforestation rates. Lowering transport costs via paved highways or unpaved roads clearly increases deforestation in the census tract in which the road investment was made. Further, at no level of prior clearing[5] did new roads decrease clearing; in fact, the largest effects of new paved highways are in heavily cleared areas (50-75% deforested). Such analyses will be done for each of the time periods, separate from pooled analyses.

This work will produce multiple papers, including for comparison of the results across periods in examining structural change over time in the deforestation processes. Those analyses and papers will be completed by the end of Year 1 of this research.

<u>Single Period, Impacts of Road Investments in Neighboring Locations</u>

For the 1976 – 1987 period, the above leaves open the possibility that a road investment within a county raises deforestation in the census tract receiving the road investment but lowers it in other tracts within that county. Again only for this first period do we have any empirical results to this point. We finds that in nearby census tracts without roads, and overall in the county, deforestation increases though in some tracts it can decrease.

This work too will produce multiple papers, again including for comparison of the results across periods in examining structural change over time in deforestation process. And these analyses and papers too will be completed by the end of Year 1 of the research.

<u>Multiple Period, Impacts of Road Investments on Deforestation</u>

The analyses above should also be pursued through joining together the data from the different periods over time. We are just starting to carry out such empirical analyses. Connecting the data across all of the periods permits further examination of dynamics (above the use of lagged variables is motivated primarily by improving the estimation). For instance, linking across periods permits analysis of the impacts on deforestation of older road investments, conditional on newer ones (and vice versa, i.e. old as controls). Such results suggest a dynamic of development facilitated by infrastructure investments, which is further investigated by linking to past clearing and other indicators of activity.

This work will be completed by the midpoint of Year 2 of the proposed research. On its own and in combination with the analyses above, it will result in multiple papers.

<u>Multiple Period, Impacts of Prior Conditions on Road Investments</u>

These data also permit empirical analysis of the correlates, and perhaps determinants, of where road investments are made. Understanding this dynamic affects all of the above. Most specifically, understanding how investments are allocated can suggest how the estimation of the roads' impacts can be improved even over the best analyses above. Certainly that is easy to speculate about in principle but empirical examination of the interactions over time between road investments and the decisions that they influence could significantly contribute to a better understanding of effects of infrastructure and, applied in a forward-looking sense, to a better sense of how planning could be applied.

Again, this work will itself result in multiple papers, again including comparison with the prior steps. Of course, all papers will be presented in seminars at various stages. This set of deforestation analyses will be completed by the end of Year 2 of the research.

[5] The number of census tracts observations was sufficient to break up the sample into four categories based on the percent of forest cleared prior to the deforestation being explains: 0 %; 0-50 %; 50-75 %; 75-100 %.

Approach to Deforestation Analyses

We model deforestation as the result of the decision by the land user on hectare *j*, who is risk neutral, about when to deforest. He selects *T*, the time when land is cleared, to maximize the expected present discounted value of returns from use of hectare *j*:

$$\text{Max}_T \int_0^T S_{jt}\, e^{-rt}\, dt + \int_T^\infty R_{jt}\, e^{-rt}\, dt - C_T\, e^{-rt} \qquad (1)$$

where:

S_{jt} = expected return to forest uses of the land

R_{jt} = expected return to non-forest land uses

C_T = cost of clearing net of obtainable timber value, including lost option value

r = the interest rate

Two conditions are necessary for clearing to occur at time *t*. Clearing must be profitable and, even if it is, (2) must hold since it could be more profitable to wait and clear at *t+1*:

$$R_{jt} - S_{jt} - r_t\, C_t + \frac{dC_T}{dt} > 0 \qquad (2)$$

If a second-order condition holds, this necessary condition (2) is sufficient for clearing.[6]

Consistent with this model, we assume irreversibilities, as trees take time to grow and development changes the marginal returns to land use. Thus empirically, we examine clearing (or non-clearing) of the standing forest, not assuming reforestation on previously cleared land. This contrasts with the analyses of stock of forest common in the literature.

In the model, deforestation occur when condition (2) is satisfied for the first time. When this occurs differs across space due to differences in land quality, access to market and both exogenous and endogenous temporal shifts. The parcel transitions determine the aggregate patterns of deforestation. We observe not parcels but rates of loss within larger areas. Aggregating the model's predictions for these areas yields our empirical approach.

Clearly we do not observe the variables in (2) perfectly within our aggregate data. As our data for the drivers of deforestation are for larger areas, our X_{it} (*i* = area, *t* = time) yield a single estimated net benefit of clearing for an entire area. The actual returns and changes in costs, though, clearly vary across parcels in that district. We acknowledge that we do not measure the parcel net clearing benefits perfectly, such that clearing occurs if:

$$R_{ijt} - S_{ijt} - r_t\, C_t + \frac{dC_T}{dt} = X_{it}\beta - \varepsilon_{ijt} > 0 \qquad (3)$$

where *i* refers to an area, *j* to a parcel, *ij* to a parcel *j* known to be in area *i*, and ε_{ijt} is a parcel-year-specific term for the unobserved relative returns to forested land uses, so:

$$\textit{Probability (satisfy (3) so that clear if currently in forest)} = \textit{Prob}\ (\varepsilon_{ijt} \le X_{it}\beta) \qquad (4)$$

Since X_{it} are the same for each parcel in an area, predictions are effectively for the larger areas' rates of deforestation during a given time interval. These predicted clearing rates depend on the X_{it} as well as on the assumed distribution of the ε_{ijt}. If the cumulative distribution of the errors ε_{ijt} is logistic, then we will have a logit model for each parcel:

$$F(X_{ijt}\beta) = (1 / (1+\exp(X_{ijt}\beta)) \qquad (5)$$

For our data, we estimate this model using the minimum logit chi-square method known as "grouped logit". If h_{it} is an area's measured rate of forest loss, then we will estimate:

$$\text{Log}\ (h_{it} / (1- h_{it})) = X_{it}\,\beta + \mu_{it} \qquad (6)$$

The variance of the μ_{it} (referring to areas) can be estimated by $(1 / (I_{it}\, h_{it}\, (1- h_{it})))$, where I_{it} is the number of forested parcels in area *i* at the beginning of interval *t*. The estimator is consistent and asymptotically normal.[7] This is estimated by weighted least squares.

---

[6] It may be violated if environmental protection becomes more stringent and the returns to ecotourism rise. Our reduced form empirics can also be interpreted in terms of both (2) and the profit condition holding.

[7] Maddala 1983, p. 30. See also an explicit discussion of heteroskedasticity in, e.g., Green 1990.

Data for Deforestation Analyses

To estimate the spatial spillovers from road investments in a given census tract, we join census-tract data on roads over time with satellite data on deforestation over time, while controlling for the effects of other observable factors that influence forest clearing.

Early deforestation maps were produced in 1997 by IBGE (Instituto Brasileiro de Geografia e Estatistica) within their Diagnostico Ambiental da Amazonia Legal product. The pre-1976 clearing is from the RADAM Project vegetation maps, with classes of land cover. The clearing in 1987 is from IBAMA/INPE maps based upon Landsat imagery. This source also provides information on the forest clearing between 1987 and 1991.

Moving forward in time (in part overlapping with the above), the TRFIC data for 1986, 1992 and 1996 provide multiple years generated by the same approach. Clearly the 1996 provides another point in time, a mid-point in the 90s decade, when that is useful. Finally the PRODES data for 2000 and 2004 will also eventually be included with above.

We tracked the evolution of roads within each census tract. The digital road maps were developed in the Department of Geography at Michigan State University from paper maps produced by DNER (Departamento Nacional de Estradas de Rodagem), all with the assistance of the Universidade Federal Fluminense and IPEA/Rio. The periods for which road-investment data will certainly be used are 1968-1975, 1975-1987, and 1987-1993.

The roads-clearing relationship is estimated in regressions with control variables, i.e. other factors expected to affect the net benefits from, and thus pace of, deforestation. These include ecological characteristics (rain, soil quality, slope, and distances to rivers) as well as distances to cities. All of these controls have been shown in previous research to have significant effects on deforestation, something that we also can confirm here.

The increase in observations, from under 300 to over 6000, that is made possible by moving from county to census-tract level permits another set of key control variables. We clearly do not observe all of the factors that drive local clearing rates across the basin. However, our detailed data permit the inclusion of spatial fixed effects for each county.

We attempt to reduce undue influence of cities by dropping about 100 census tracts near to 19 large cities (i.e., in 20km of a city with density over 100 people/km^2). As a robustness check, we also drop about 1000 tracts near 270 medium and large cities (i.e., in 20km of a city with density over 11 people/km^2). This appears within our tables.

Overall Timeline

As suggested within the overview of the types of deforestation analyses proposed above, those foci on total deforestation have a two-year timeline for producing analyses/papers. During the third year of the proposed period, dissemination and comparison with other efforts on this important question (which is always ongoing) will be one major activity, while for new research this approach will be extended to consider forest fragmentation.