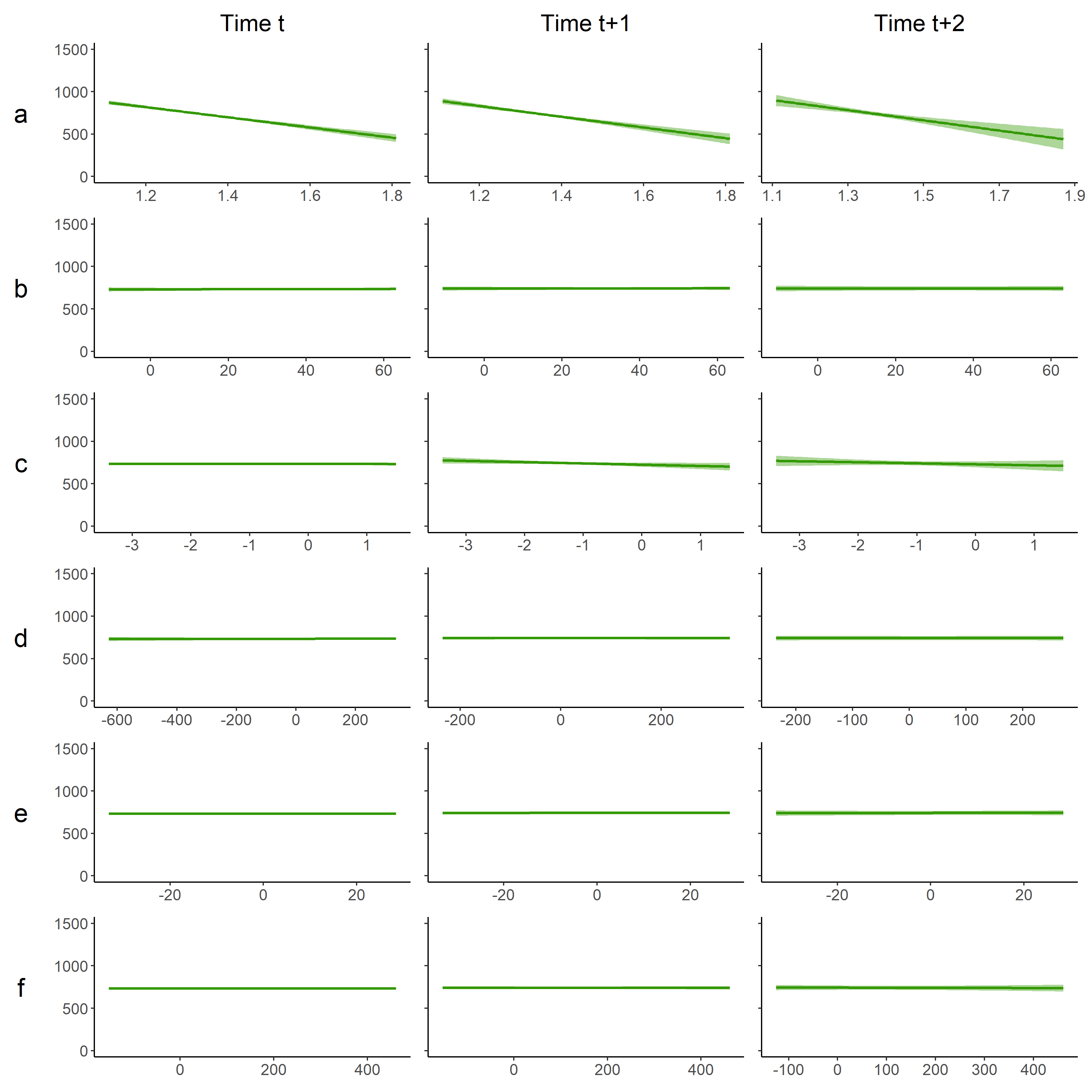
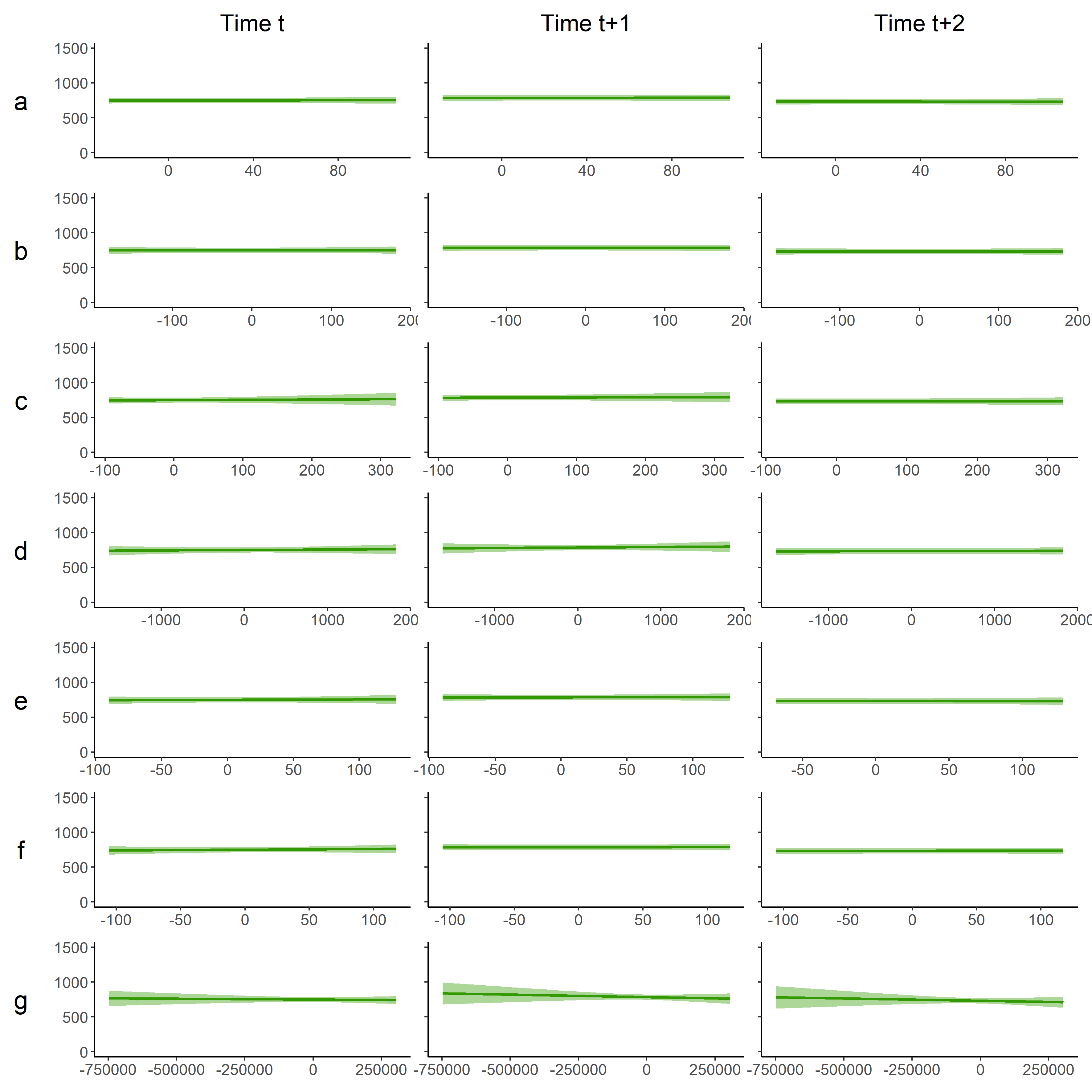
Chapter 1 results summary

**Macroeconomic predictors of forest cover change**

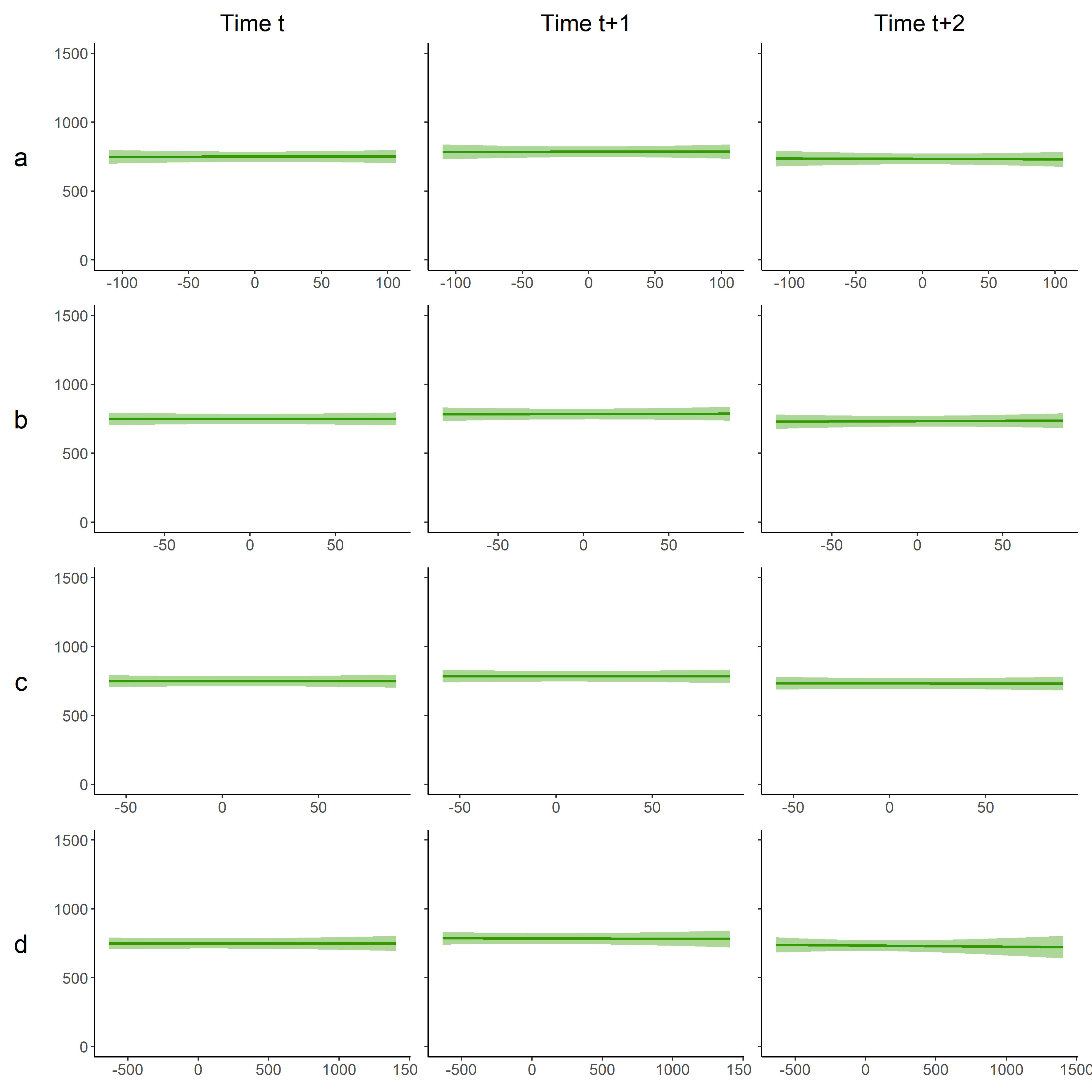
*Set 1 – macroeconomics*



**Figure 1. Predicted relationship between forest loss and macroeconomic variables. All y-axes are the amount of forest lost in hectares. Row a: population density (individuals/km2), row b: Gross Domestic Product (GDP), row c: agricultural sectors contribution (%) to GDP, row d: development flows to the agricultural sector (USD millions), row e: development flows to the environment sector (USD millions), row f: Foreign Direct Investment (USD millions).The left column of plots are the effects on forest cover at time t (i.e. the variable values and forest loss values from the same year), the middle column of plots are the effects at time t+1 (i.e. the effects on forest loss in the subsequent year), and the right column of plots are the effects at time t+2 (i.e. the effects on forest loss two years after the variable values).**



**Figure 2.** **Predicted relationship between forest loss and commodity variables. All y-axes are the amount of forest lost in hectares. Row a: Crop Production Index, row b: Non-food Production Index, row c: median annual market price for rice (USD/t), row d: median annual market price for rubber (USD/t), row e: median annual market price for corn (USD/t), row f: median annual market price for sugar (USD/t), row g: total production from forestry (m3). The left column of plots are the effects on forest cover at time t (i.e. the variable values and forest loss values from the same year), the middle column of plots are the effects at time t+1 (i.e. the effects on forest loss in the subsequent year), and the right column of plots are the effects at time t+2 (i.e. the effects on forest loss two years after the variable values).**



**Figure 3.** **Predicted relationship between forest loss and the producer prices (i.e. farmgate prices) variables. All y-axes are the amount of forest lost in hectares. Row a: producer price for rubber (USD/t) row b: producer price for cassava (USD/t), row c: producer price for corn (USD/t), row d: producer price for sugar (USD/t). Left column of plots are the effects on forest cover at time t (i.e. the variable values and forest loss values from the same year), the middle column of plots are the effects at time t+1 (i.e. the effects on forest loss in the subsequent year), and the right column of plots are the effects at time t+2 (i.e. the effects on forest loss two years after the variable values).**

**Socioeconomic predictors of forest cover**

This section of the analysis uses a national socioeconomic survey called the Commune Database. I have data from 2007 to 2012. The purpose of this analysis was to model the relationships between socioeconomic factors and forest cover to see if forest cover could be predicted by socioeconomics. The dataset was originally at the village level, but I aggregated up to the next administrative level – the commune.

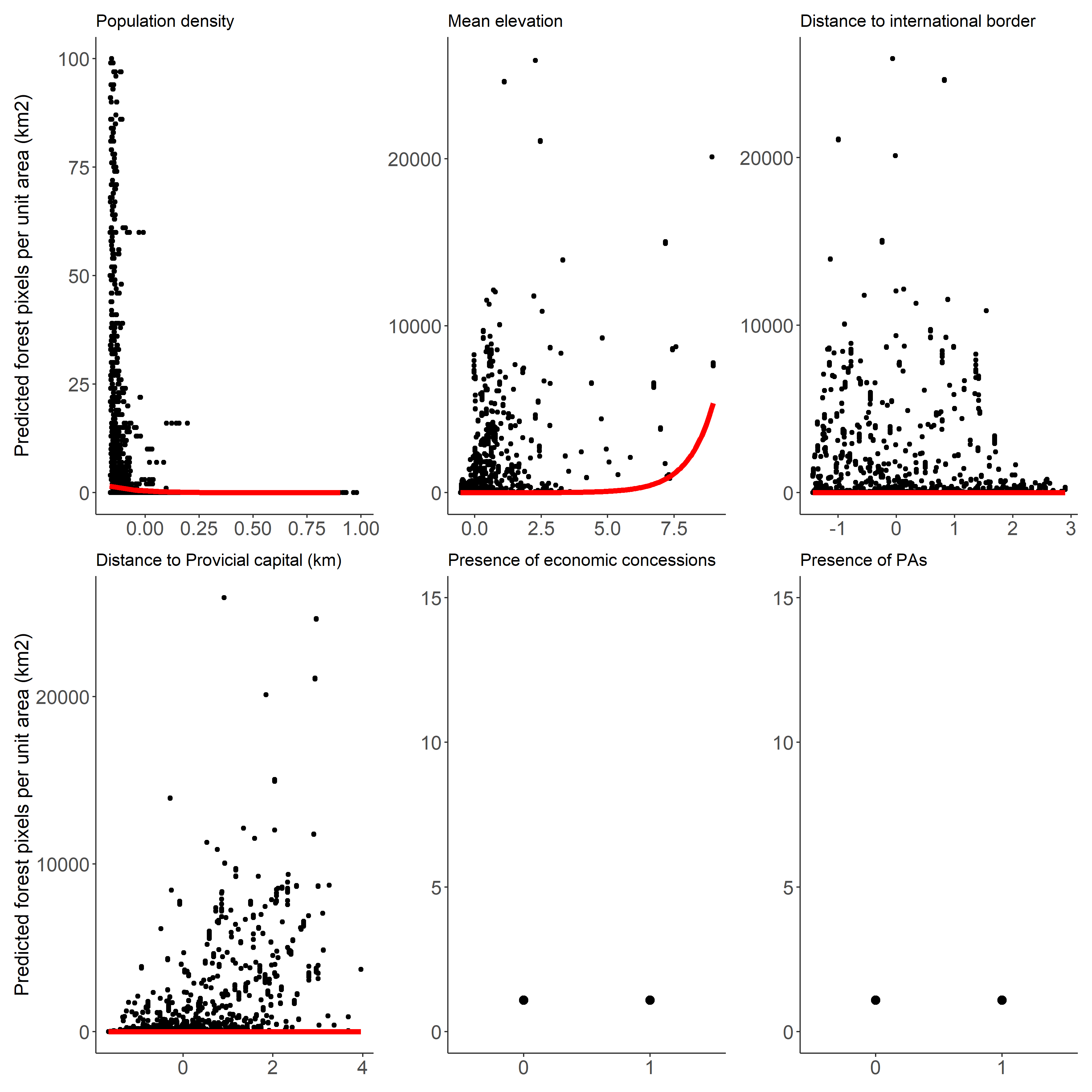
I had 21 predictor variables in total, split into 9 sets. Because I was not explicitly interested in changes over time, I used generalised linear mixed models, with year as a random effect. I had commune as a random effect to account for repeated measurements, and commune was nested within Province, as I was not explicitly interested in the effect of province, but I was interested in how the other predictor effects might vary between provinces.

First, I fit GLMMs to each of the predictor sets to explore potential effects. These models highlighted that there were very few effects of note. I used these results to select the predictors with the strongest (relatively) effects, and these were carried forward to a global model. I then built a full global model and removed variables stepwise until I had a final model.

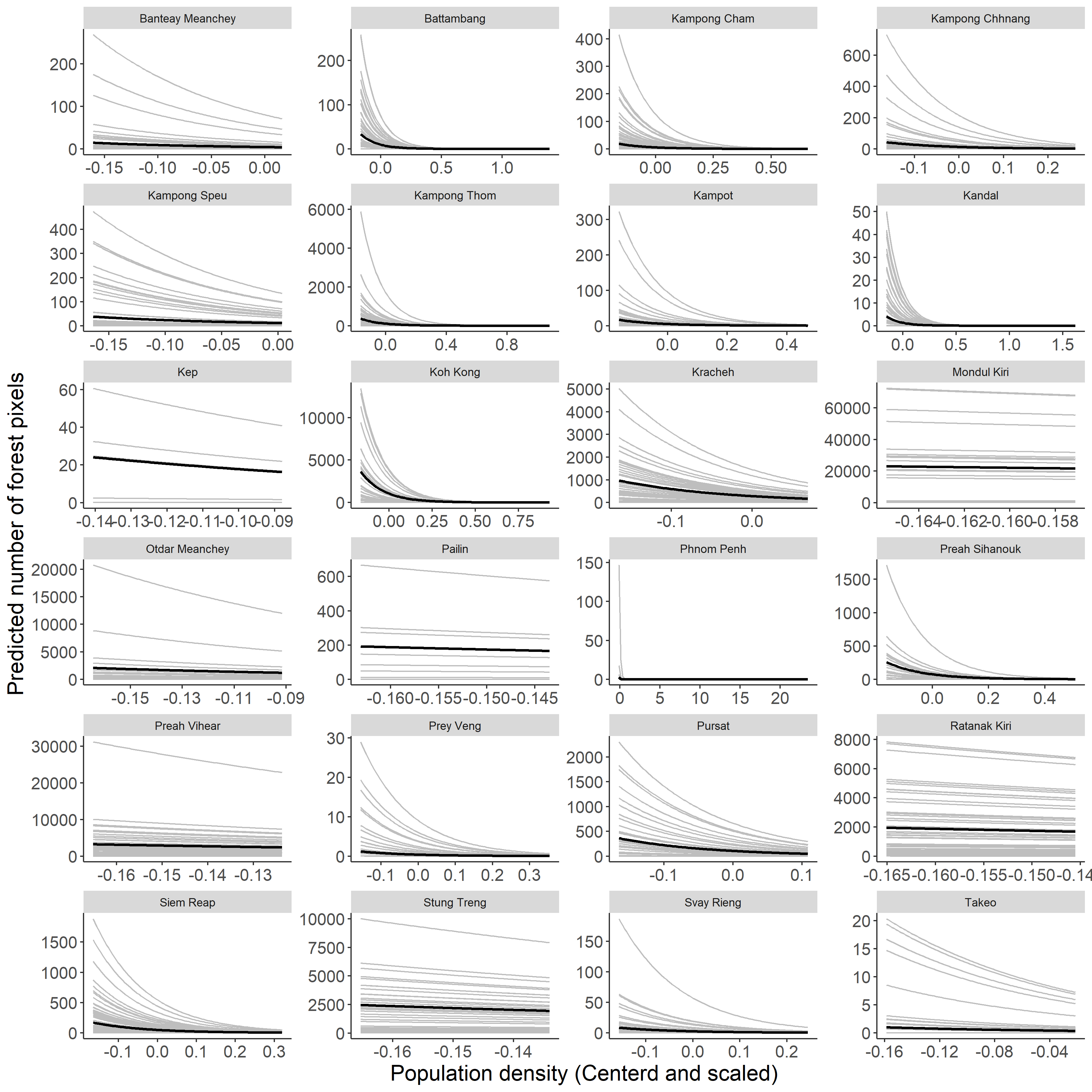
I ran global predictions from the final model (Fig 4), and I also ran predictions between provinces (Fig 5 & 6). For the between-province predictions, I predicted for each commune within a province, and then took the 50% quantile as the “mean” provincial prediction and used the 2.5 and 97.5% quantiles as confidence intervals, or “variance intervals” to display the within-province variation.

As there were serious issues with the above global model, and the resulting predictions (see Fig 4), I further aggregated the data up to the provincial level. This was intended to neutralise the between-commune variation. Here the random effects structure was year as a random slope and Province as a random intercept. First, I ran GLMMs with the variables as continuous as I did above, but the raw data suggested that there were two “peaks” in the data for some of the variables, and the resulting predictions were clearly very poor (Fig 7). The shape of the data suggested there were potentially two “types” of province for some of the variables, and so I converted the data into “high” and “low” categories by splitting the variables by their mean. These models again returned very little of interest (Fig 8). The only significant effects were from economic land concessions and protected areas, where the model predicted significantly higher forest cover in provinces where there were land concessions and protected areas. This was not surprising, as the provinces with very little forest cover are unlikely to have PAs. It is quite interesting (although not at all surprising) that provinces with high forest cover are more likely to have economic concessions, as it is well known that economic concession (i.e. industrial agricultural concessions) were almost exclusively awarded on forested land. This allowed the concession companies to make large profits from timber extraction prior to planting their agricultural crop.

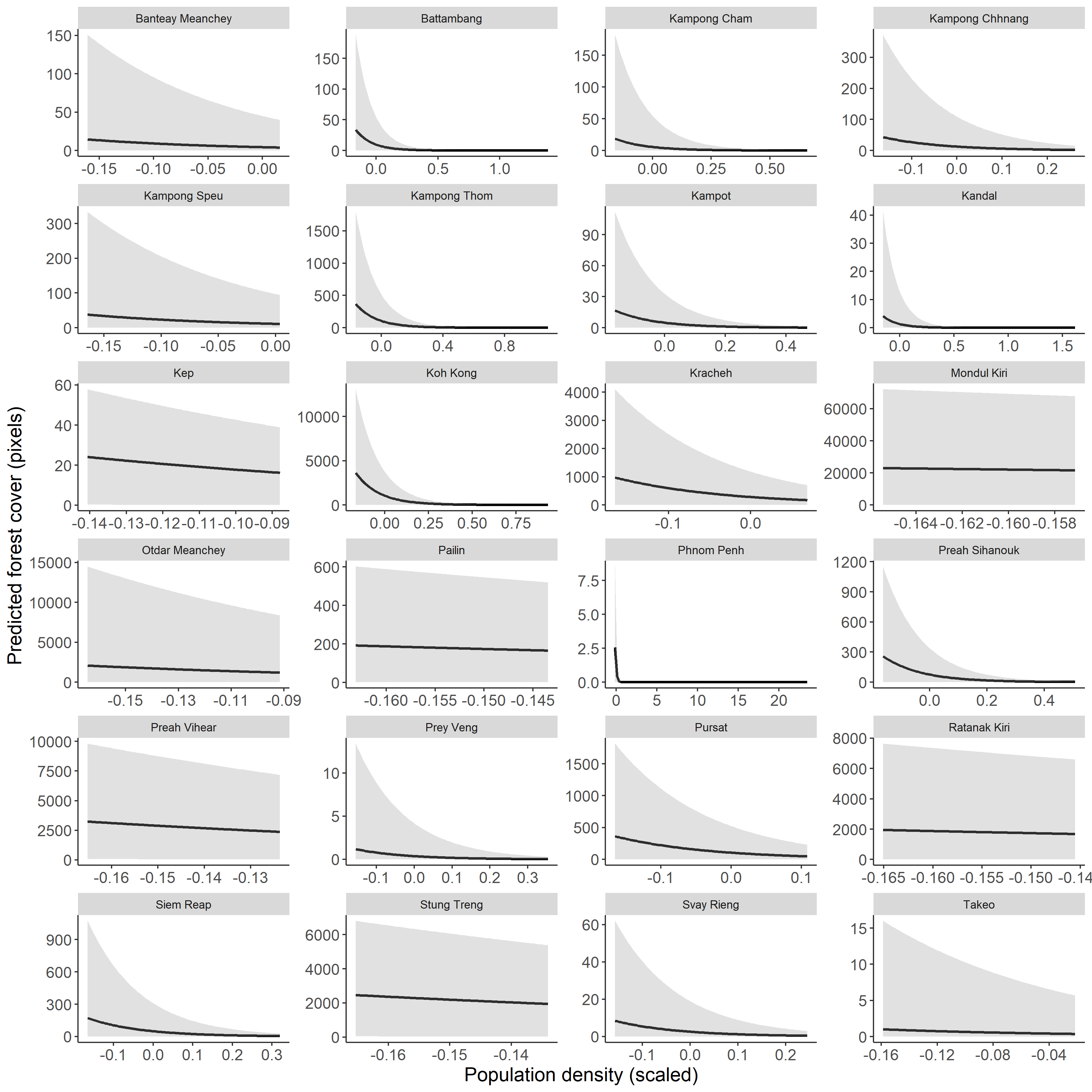
Finally, in the absence of any particularly interesting results I conducted a cluster analysis to see if the provinces could be categorised into broader typologies based on socioeconomics (and environmental factors, including forest cover).



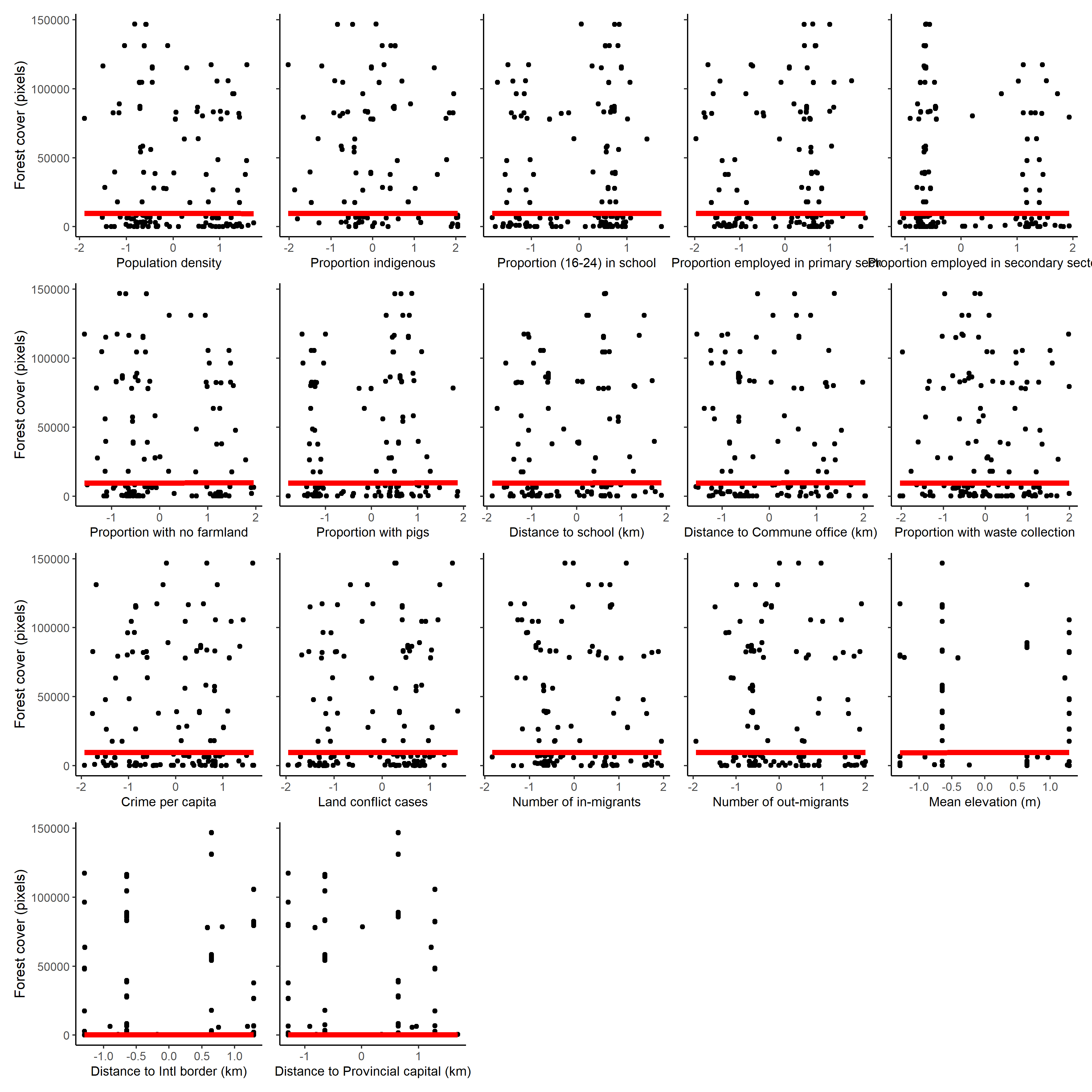
**Figure 4. Global predictions (i.e. predictions for an average commune within an average province) using the final model. X axes are the scaled predictors. Plots show the predicted values (red line) over the raw data.**



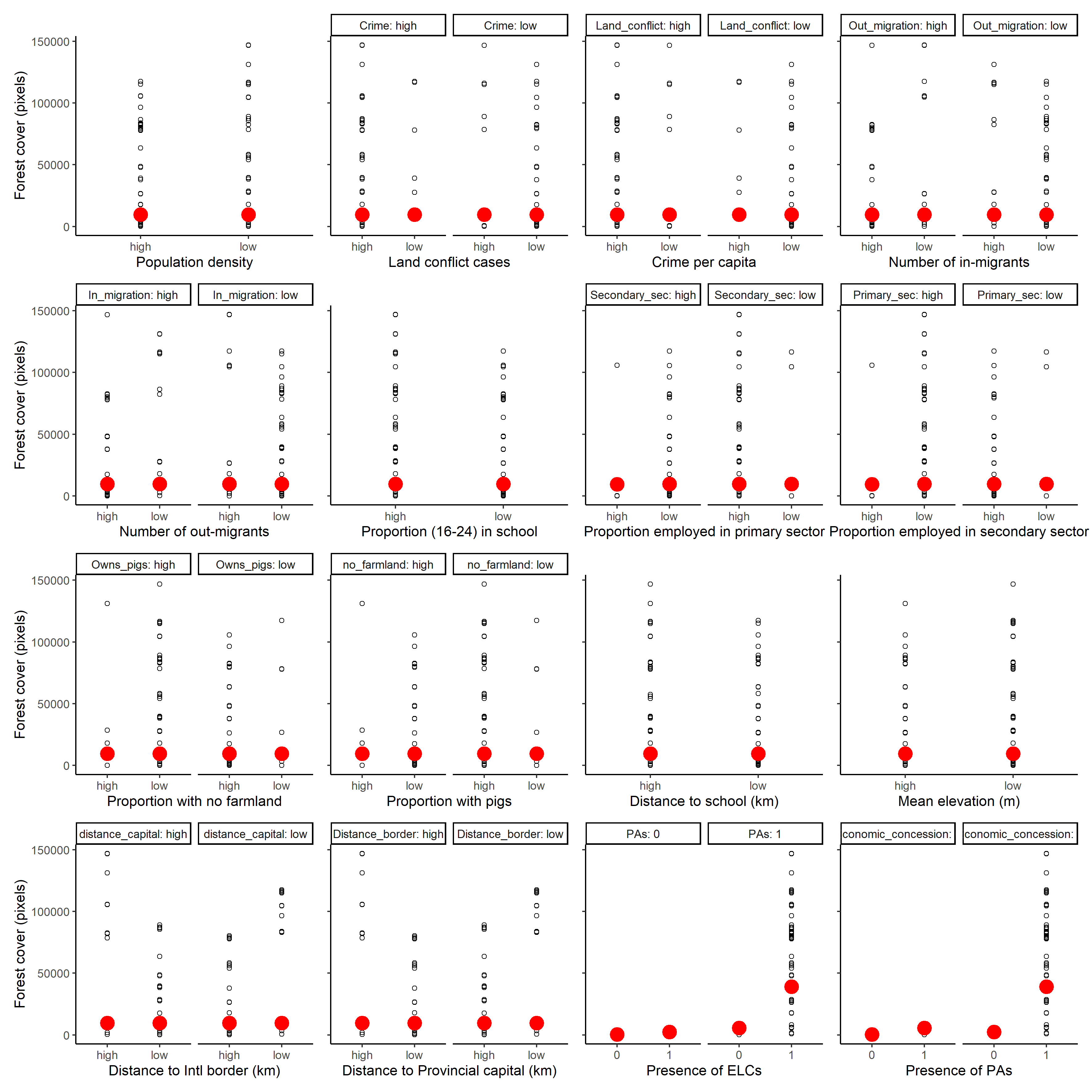
**Figure 5. Predictions within each province using the global model. Thick lines are the mean provincial predictions which are calculated as the 50% quantiles from all the commune-level predictions. Faded grey lines are the individual commune-level predictions.**



**Figure 6. As above - predictions within each province using the global model. Thick lines are the mean provincial predictions which are calculated as the 50% quantiles from all the commune-level predictions. Error ribbons are the 2.5% and 97.5% quantiles from all the commune-level predictions within each province.**



**Figure 7. Predictions from model sets using provincial-level data. Red lines are predictions, black points are raw data.**



**Figure 8. Predictions from model sets using province-level data. Red points are predictions, black points are raw data. For model sets with more than one variable, the plots have been faceted to show predictions from all combinations of variable levels.**