Labor Market Impacts of Bolivia's Protected Areas

Vanessa Echeverri Figueroa

Advisor: Gustavo Canavire Bacarreza

Universidad EAFIT, Medellin Colombia

June 8, 2016

Abstract

Protected areas are a powerful policy instrument in the preservation of the ecosystem and global biodiversity. However, measurable socioeconomic effects such as poverty or labor market effects are still not well understood. Some recent studies show evidence of heterogenous poverty reduction effects, but there is no compelling evidence on the labor market side despite the fact that poverty is typically battled through structural changes in the labor market. By employing non-parametric techniques we find evidence that supports that instituting protected areas has positive effects on labor markets. Despite the indisputable benefits that are obtained by the increased preservation of fauna and flora for the country, there is consistently evidence linking the latter policy decision with a slight reduction in unemployment, and a decrease in informality as well as in jobs in extractive and natural resources industries.

Keywords: Protected Areas, Labor Market, Unemployment, Informality, Bolivia

Introduction

The establishment of protected areas, or protected areas, is popular for being a powerful policy tool to obtain an instant preservation of a determined ecosystem; in turn rendering its biodiversity a national treasure. The sheer coverage of protected areas worldwide makes for an important subject of investigation aiming at understanding the underlying socioeconomic impacts associated with such policy decisions. However, there is unconvincing evidence to determine the effects and channels through which protected areas impact socioeconomic indicators, such as poverty, unemployment, and labor migration. Canavire-Bacarreza & Hanauer (2013) and Ferraro et al. (2013) found evidence corroborating the hypothesis that protected areas are a key element in the reduction of poverty in Bolivia; however, little evidence is found regarding the origins and channels through which socioeconomic variables are struck. The lack of empirical evidence for such hypothesis is a fundamental setback for resolving the internationally pertinent debate of; whether the establishment of protected areas counteract poverty reduction goals in developing nations and if these can have an effect on the labor market structure?

Regardless of the uncertain effect of protected areas, there were significant advances in the methods and theory to explain such effect. Most studies tend to solely focus, either, on the impacts of protected areas on socioeconomic outcomes (Andam et al. (2010),Canavire-Bacarreza & Hanauer (2013), Brockington et al. (2006), Bandyopadhyay & Tembo (2010)) or on the environmental outcomes (Andam et al. (2008), Andam et al. (2013)). There are few studies that examine the joint environmental and socioeconomic impacts of protected areas (exceptions include Sims (2010)) and just a hand-full aim at explaining the channels in which the impact travels on and the effect that such policy decisions have on the labor market (Ferraro & Hanauer (2011),Ferraro et al. (2011)).

In this study quasi-experimental methods are implemented in order to examine the labor market impacts of Bolivia's protected areas. We estimate the impacts that Bolivia's protected areas have on labor market outcomes such as unemployment, informality, and the concentration of natural resource extraction jobs between 1992 and 2012. Protected areas were first established in 1992 (when protected areas were first recognized by common law (see Canavire-Bacarreza & Hanauer (2013)) and the country data reflects that they continued to grow in number until 2012. This study provides several significant contributions for the purpose of further understanding the socioeconomic bearings of Bolivia's protected areas. In regards to the empirical methods utilized throughout this investigation into labor market repercussions due to natural conservation policy, significant evidence of a causal relationship was also found. With a strict scope in Bolivia's protected areas and the country as a whole, we arrive at two fundamental findings; the first is derived from an extension of Canavire-

Bacarreza & Hanauer (2013) where we make use of an updated census data (2012 as opposed to 2001) and a more disaggregated geospatial socioeconomic data set. This modification allows for an increased level precision and more relevant estimates of Bolivia's environmental sectioning off impact on poverty and labor market effects. Secondly, we utilize multiple estimation strategies to compare the impact, measured through both methods, of Bolivia's protected areas on its labor market.

Data

Previous studies indicated that the establishment of Bolivia's protected areas resulted in reductions in deforestation Ferraro et al. (2013) and reduced poverty Canavire-Bacarreza & Hanauer (2013) . To estimate the labor market effects of protected areas we draw from these previous studies and employ two distinct data sets: the deforestation data and the socioeconomic data. All geographic information systems (GIS) calculations in ArcMap 10.x, (Qgis and SagaGIS that is tool of Qgis) was conducted for the analysis.

Unit of Analysis and treatment assignment

Bolivia's political divisions are composed of nine (9) departments or states, one-hundred and twelve (112) provinces, three-hundred and twenty-seven (327) municipalities and one-thousand three-hundred and eighty-four (1,384) cantons. Due to the "Ley Marco de Autonomías de 2010," cantons are removed from the map. An analysis at the canton level was done (at the political subdivision of the municipalities), thus they are smaller than a municipality but larger than a community. Few cantons were found that are as large as a municipality and that are as small as a community. There is Information for one-thousand two-hundred and fifty-two (1,252) cantons, of which one-hundred and six (106) are protected, one-thousand and seven (1,107) remain unprotected. Protected units run according to Canavire-Bacarreza & Hanauer (2013) and follow the guidelines established at the 4th World Congress on National Parks and Protected Areas. They concur with the common measurement criterion that if ten percent (10%) or more of the total area of the unit is occupied by one or more protected areas, it is deemed a "protected area" or "natural reserve." A "slightly protected" unit has an area that is occupied by a protected area between [0.01, 0.1]. According to the previously mentioned premises, it is necessary to remove 39 slightly protected cantons from the sample to reduce the potential contaminated controls, thus obtaining a higher level of precision.

Biophysical Data

Geographic Information Systems (GIS and QGIS) are used to distinguish and analyze between three central categories of spatial data; 1) key biophysical qualities that affect the assembly of protected areas and unemployment, 2) territorial mappings of municipalities for the 1992, 2001, and 2012 censuses in addition to the veiled demographics, and 3) temporally distinct boundary mappings of terrestrially protected areas. Roadmaps and other geographical data, such as elevation, flora density, metropolitan centers, among other crucial information, was obtained from National Aeronautics and Space Agency (NASA), conservation International, and the Bolivian Forest Regulation Office (or *Superintendencia Forestal* in Spanish). Natural reserve boundaries were provided by the National Service of Protected Areas (or *Servicio Nacional de Areas Protegidas* in Spanish) (SERNAP), the World Database on Protected Areas (WDPA).

Socioeconomic Data

Three recent censuses of Bolivia's population and housing conducted in 1992, 2001 and 2012 were obtained from Bolivia's National Statistics Institute (INE). The data captures socioeconomic and employment variables for the three years. The three censuses, serve in estimating the socioeconomic indicators used in this investigation, such as unemployment, poverty, education, housing and health, in a canton scope.

Methods

Matching

As in previous studies Canavire-Bacarreza & Hanauer (2013) the analysis ATT (average treatment effect on the treated) was used to measure the impact of protected areas on Bolivia's labor market, however, to better understand why it is relevant to implement the ATT it is important to begin with a background in matching.

There exists a fundamental problem in the measurement of the difference found between the outcome variable and the voluntary subject present in the program treatment effect, or program. The issue in measuring the impact lies within the outcome variable of the subject in the absence benefiting from the program itself. In order to construct a model that adequately represents the effect of treatment (given the fundamental problem) through its representation in impact studies, lies in identifying the difference between the outcome variable of the individual participant once he or she has implemented the program and the outcome variable that the same individual would have obtained in the hypothetical case that the program was unavailable, or non-existent. Because of the hypothetical case that the program is nonexistent, the observations of this result are omitted; these are also known as the counterfactual results. It is impossible to view both results for the same individual simultaneously, given the binary assembly of the hypothesis.

In the case of a binary program the treatment indicator (D_i) would be equal to one ($D_i = 1$) if he is receiving treatment or the treatment indicator (D_i) would be equal to zero ($D_i = 0$) if there exists exclusion of the individual from the program. Each individual's outcome variables are defined as $Y_i(D_i)$, for each individual $i = 1...N$. In a binary case it would be taken into account one result per subject; either, $Y_i(1)$ if the individual is treated $D_i = 1$ or $Y_i(0)$ if the individual is left untreated $D_i = 0$. The resulting effect of treatment of an individual i, or impact of the program can be described as follows:

$$\tau_i = \gamma_i(1) - \gamma_i(0) \tag{1}$$

The formation of a binary system results from the given parameters that the individual can only opt to be treated or not. The impact of the program measured by the previous equation 1, refers to a given period of time when the decision is made, hence, it does not allow us to compare the subject in two different instances within a given timeframe. The period of time under study is divided into two phases; the pre-treatment phase or the moment that the subject makes the decision and the second phase constitutes the time period after the individuals are treated, post-treatment, hence, the observable result can be described as follows:

$$\gamma_i = D_i\gamma_i(1) + (1 - D_i)\gamma_i(0) = \begin{cases} \gamma(1)_i & if \quad D_i = 1 \\ \gamma(0)_i & if \quad D_i = 0 \end{cases} \tag{2}$$

If the necessary information is present, and it has been registered into the resulting data, the variable is understood and recorded as an observed variable. Conversely, if the variable was unobserved, or is unrecorded in the data set, it is considered an *unobserved variable*.

Given that the unobserved variable for each subject, i, is not estimated if he or she opts out of the treatment, $D_i = 0$, it is impossible to estimate the treatment, τ_i. Subsequently, it is recommended that one should focus on the permeation of average effect that the program has on the population, or focus groups of the population, under analysis.

In order to quantify the *average treatment effect* (ATE) of the impact on the population, yielded by the application of the program, the following equation is used:

$$\tau_{ATE} = E[\gamma_i(1) - \gamma_i(0)] \tag{3}$$

Where $E[\cdot]$ represents the operator of expectations.

When studying the impact of a program from a universal scope, it is relevant and necessary to assume the representation of average effects that variables have on the population. When an individual, who is randomly chosen, is elected to take part in the program, the average change of the outcome variable yields an effect of ATE. However, in reality there are few cases in which a universal program can be implemented on the population as a whole; hence, a focused sample of the population is frequently implemented to represent the general tendencies and qualities of an entire group. While analyzing the representation of the population based on a sample, it is advantageous to utilize the ability of a unique estimator that only averages the effect of the eligible population.

The average measurement of treatment of the treated (ATT) provides critical insight concerning the difference between the mean of the outcome variable and the mean results of subjects corresponding to the focus group or sample population under analysis who, otherwise, would not have received any treatment. Given set of individuals, the ATT forms the chief parameter within these types of studies that center on evaluating the impact that the programs have. The ATT is represented by the following equation:

$$\tau_{ATT} = E(\tau_i \mid D_i = 1) = E[\gamma_i(1) \mid D_i = 1] - E[\gamma_i(0) \mid D_i = 1] \tag{4}$$

Where $E[\cdot]$ represents the operator of conditional expectations.

If $E[\gamma_i(1) \mid D_i = 1]$ represents the expected value of the outcome variable representing the group undergoing treatment, and $E[\gamma_i(0) \mid Di = 1]$ would be known as the counterfactual result or expected values of the outcome variable of those groups that are not treated. The counterfactual average of those individuals who could have been treated but were non-existent, or unobserved, within the group, constitute a hypothetical result that would not be otherwise recorded in the dataset or observed in reality. The relevance of the average effect is particularly useful in order to define and determine the continuance of a running program or its elimination.

Selection bias

The principal measurement challenge lies in determining the conditions in which $E[\gamma_i(0) \mid D_i = 0]$ can be undertaken as a valid approximation of $E[\gamma_i(0) \mid D_i = 1]$ and if it can be used in equation (4) to estimate the effect of the program. It is notable that $E[\gamma_i(0) \mid D_i = 0, X]$ can be effected as a trustworthy approximation of the counterfactual if $E[\gamma_i(0) \mid D_i = 0, X] = E[\gamma_i(0) \mid D_i = 1, X]$ given a set of features that jointly affect the establishment of the program. The previous expression means that the outcome variable in the absence of the program is equal to the group of treated individuals ($D = 1$), than for the groups of individuals with a nonexistent treatment ($D = 0$).

Every time an individual voluntarily elects to participate in the treatment program, the assumption $E[\gamma_i(0) \mid D_i = 0, X] = E[\gamma_i(0) \mid D_i = 1, X]$ is violated. The previous assumption is violated when each individual is unique and they voluntary decide if they wish to be treated or if they opt out of the treatment even though they are eligible to receive the treatment. There is a presence of observed and unobserved characteristics that differ between individuals which is probably a driving factor in the differing outcome variable of each group, treated and non-treated. Even if the program is nonexistent, this differentiation is a likely cause of divergent outcome variables for each group. This effect is known as the selection bias.

τ_{ATT} can be written as:

$$E[\gamma_i(0) \mid D_i = 1] - E[\gamma_i(0) \mid D_i = 0] = \tau_{ATT} + E[\gamma_i(0) \mid D_i = 1] - E[\gamma_i(0) \mid D_i = 0] \quad (5)$$

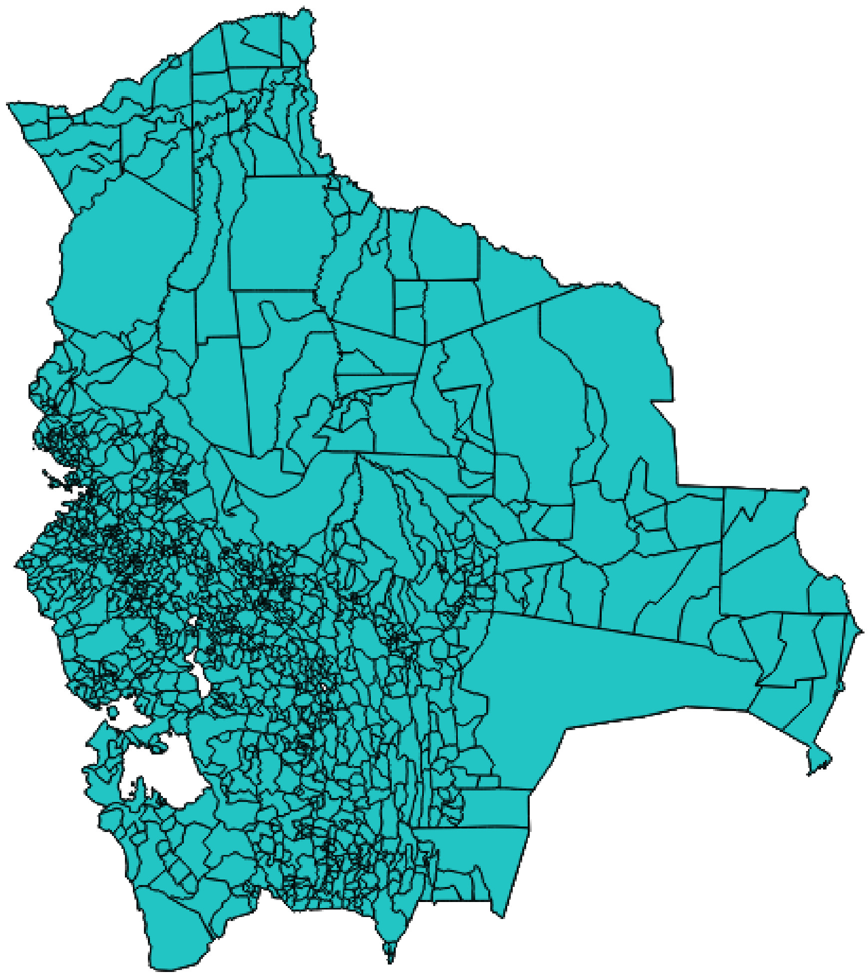
The previous equation represents the measurable difference found between individuals of the treatment group (represented on the right side of the equation) and the control group (the equation is valid even if the program is non-existent). Hence, the average of the control group (represented in the left side of the equation) will be equal to τ_{ATT} plus the pre-existent difference between both groups is the biased estimator. The direct effect of treatment is derived from the difference of the averages between both groups. Without any additional information, it is impossible to determine the source of each measure.

Construction of the Canton map

Since this map is currently unavailable, Bolivia's canton map has been built based on shape-files of points and municipalities, where the shape-file of points shows the center point of each locality in Bolivia during 2001. The municipalities' shape-file shows the three-hundred and twenty-nine (329) municipalities within Bolivia's map. Each municipality contains, within its boarder, several cantons that are permanently subscribed to each municipality. The shape-file of points contains each (the) localities' codes so it is possible to identify its parental canton. Codes are assigned depending on the location, as the first two positions (digits) make reference to the nine departments of Bolivia, third and fourth positions (digits) make reference to provinces, fifth and sixth make reference to the municipalities, the seventh and eighth make reference to the cantons, and the ninth and tenth reference the localities. A series of Thiessen's polygons have been created subscribed to each municipality, so at first instance there was the localities' map but it was necessary to dissolve those new polygons, which had the same canton code, so that the borders of those cantons that fall within the characteristic scope of this investigation could be created. Thiessen's Polygons are those built around points such that each area within the polygon is closer to a point inside that particular polygon than any other point in the sample (Wade et al., 2006).

When Thiessen's polygons were generated, it was evident that these were not subjected to the municipality's geographical limits, so it was required to correct the boundaries so they would adequately represent each municipality. The shape-file of points has been cross-referenced with the respective boundaries of each municipality's polygon and it was necessary to defragment the municipalities' map into single polygons. The latter segmentation generated Thiessen's polygons within municipalities, then they were clipped into each municipality's polygons in order to adjust both shape-files' external borders. All of the Thiessen's polygons with the same canton code have been dissolved into one polygon. The latter process was repeated for each Bolivian municipality. At the end there were three-hundred and twenty nine (329) shape-files containing canton borders and they were merged onto one layer for ease in illustration. Merging the shape-files allows for a seamless union of data to facilitate the display of the shape-files in accordance to Bolivia's unique canton outlines. These shape-files have been created under the impossibility of multipart polygons; however, the previous assumption was not undertaken for multipart polygons separated by bodies of water.

Figure 1: Constructed Bolivia's Cantons Map



During the process of segmentation, several issues arose in the process, being one of those multipart polygons problem that consists in those polygons sharing the same canton code but with different geographical coordinates. In order to neutralize the multipart polygons issue, some of them were merged with respect to their size. In order to eliminate any redundancies with each polygon's canton code, the smaller polygons have been joined with whichever neighboring polygon had an inferior area to that of the other cantons. During the analysis, each municipality polygon was separately treated. But when all layers were merged into one there were some overlapped canton polygons that overlaid their boundaries into various municipalities. This violated the assumption, so a similar criterion was applied to this problem by merging the repeated canton codes that did not correspond to its municipality code into the neighboring canton with an inferior area. The previous process was realized while aiming to maintain the singular code that had been originally assigned by the state. Cantons contained within other cantons were also merged. Attributes from the largest areas and those points outside of the range have been excluded from the resulting illustration that can be considered a comfortable approximation to the 2001 regionalization of Bolivia.

Outcomes

Three labor market indicators were chosen as outcome variables to determine the impacts of protected areas; unemployment rate, informality, and the proportion of the labor force that is dedicated to agricultural work, livestock farming, and natural resource extraction. It was decided that other labor market indicators be excluded such as, underemployment, global participation rate, age differing indicators, among others, in order maintain a narrow scope of investigation and, thus, more robust results. In accordance to the narrowed focus, it was found that the chosen variables (protected areas, unemployment rate, informality, and the proportion of farm laborers) provide the clearest connection when determining the effect of protected areas on the previously mentioned chosen indicators.

(a) Unemployment rate

According to conventional methods of segmentation of the labor market, the unemployment variable parameterized in this study was defined as the open-unemployment rate which is equal to unoccupied as a fraction of the labor force:

$$U_i = \frac{unoc_i}{lf} * 100 \tag{6}$$

Where U_i is the open unemployment rate for the canton i, $unoc_i$ are the number of unoccupied in the canton i and lf represents the labor force. The labor force includes individuals who are occupied and unoccupied and are actively seeking work or are presently employed. An occupied individual is one that has worked within one week, is committed to his or her house chores or worked, all within that one-week period. The unoccupied are those individuals who have previously worked and are currently seeking employment. The unoccupied are further segmented into those; who are seeking employment and have worked in the past (aspirant); and those who are seeking employment for the first time (unemployed).

(b)Informality

General census distinguishes the labor informality variable by segmenting into self-employed individuals plus, cooperative of production, family employed and unpaid internships; all of which are divided by the occupied segment

$$I_i = \frac{SE_i + CP_i + FUI_i}{Occupied} * 100 \tag{7}$$

Thus, I_i is the informality for the canton i, the first component of the numerator describes the self-employed individuals in the canton i, CP_i represents the cooperative of production in canton i, the third term represents the family workers and unpaid interns in canton i.

(c) Proportion of agricultural workers

In this analysis we capture those workers who dedicated their time towards naturally extractive activities such as, fishing, agriculture, livestock farming, hunting, forestry, and mining operations. The previous activities were pooled into one outcome variable. For the pooled activity, we aimed to identify the change in the proportion with respect to designated protected areas.

$$\%Agricultural_i = \frac{EA_i}{TBEA} * 100 \tag{8}$$

Where EA_i denotes all the workers in canton i dedicated to the aforementioned extractive activities and TBEA represents the number of workers in total branches of economic activities.

There was a change in the method of capturing each branch of economic activity undertaken by the Bolivian labor force, in which it is found seventeen (17) subdivisions of general activities in the 1992 and 2001 census, which were utilized in these estimations. As defined by 2012, the nomenclature of subdivisions of economic activities was further divided into 21 subdivided economic activities [1].

[1] See the full list of the economic activities in table 13, Appendix.

Moderating Covariates

In order to get a clear perspective of the driving variables that have a significant impact on the casual effect of protected areas in Bolivia, a set of observable covariates were chosen in accordance to Canavire-Bacarreza & Hanauer (2013) which are the main covariates that determined the establishment of protected areas and are necessary to measure their effect relating to labor market outcome variables. Hence, it must be assumed that the remaining unobservable variables do not cause a significant impact, or influence. For the purpose of controlling the observable characteristic that impact a unit's reaction to a given treatment are the moderating covariates. For the case are then a) Baseline poverty, b) forest coverage in 1991 (%), c) distance to major city (km), d) Average elevation (m), d) Average slope (%), and e) Roadless volume (km). Simultaneously impact of protection for both, protected and unprotected cantons was isolated.

The *poverty index* (PI) is an asset-based poverty index extracted from Bolivian household censuses conducted in 1992, 2001, 2012. The PI is a critical anchor in order to measure the impact on poverty generated by sectioning off protected areas within the state. There is a general assumption that the PI contains a number of household assets and characteristics that explain the variation in unobserved poverty outcomes. The use of factor loading (eigenvectors) measurements brings forth the relative influence of each component. The latter measurement is derived from the ratio yielded by variance/covariance matrix of the component variables. In order to measure poverty, principal component analysis (PCA) was utilized. The combination of eigenvectors and relative municipal-level variation in assets serve to compute a municipal level PI. The eigenvectors provide factor scores F_j for asset $j \in \{1, 2, ..., J\}$ which serve to identify the weight and vector of the influence each asset ai exerts on the PI. The combination of the factor scores with asset levels comprise the PI for any Canton $i \in 1, 2, ..., N$,

$$PI_i = \sum_{1}^{J} F_j \left[\frac{a_{ij} - \bar{a}_j}{s_j} \right] \tag{9}$$

a_{ij} the observed level of asset j for canton i, $\bar{a}_j$ is the average of asset j across all cantones s_j is the standard deviation of asset j across all cantons.

To assure uniformity or at least comparability between the three PI tallied, data was pooled for the PCA and estimate the mean influence of each asset within the given timeframe. Asset variation drives the estimated changes in poverty. The insight extracted from understanding how and the level of household asset co-variation provides forth the composition of these assets across cantons and relates their variation across cantons regarding their relative poverty levels. A resulting negative factor score signifies that the asset variable causes surges in poverty, and *vice versa*. The analysis yields evidence of the internal validity of the constructed PI as it increases the trends in asset levels similar to effects experienced from increases in wealth, indicating that the PI is likely capturing poverty.

Distance to Major Cities. Growth and development tend to gravitate and develope within the metropolitan areas in most countries, hence, measuring the distance to major cities proves to be a significant variable in determining the effect of implemented programs. Reserved areas tend to be located in remote areas further from metropolitan centers in order to reduce the opportunity cost associated with the appropriation of the reserved land. The average distance from each county to the nearest city was calculated using GIS (where each municipality is segmented into 1 ha parcels and the average Euclidean distance from both endpoints). The cities included in these measurements are La Paz (and El Alto), Sucre, Cochabamba, Cobija, Trinidad, Oruro, Potosí, Tarija, and Santa Cruz.

Roadless Volume. The level of infrastructure is a clear determinant of economic development and growth, therefore the Roadless Volume is a key factor when searching for insight into establishment of protected areas, the causality of those and their effect on the labor market indicators. Previous studies have supported the hypothesis that protected areas are not typically found near dense, or complex, motorway grids. Roads serve as a great indicator of infrastructure development and urbanization levels within a given region. Roadless volumes are generated through the Euclidean distance measurement between a road and natural reserve patch. The volume is calculated from the sum of the product of the area of each 1 Ha land patch and the distance from that patch to the nearest roadway.

The *Elevation and slope* of the regions under study determine the productivity and relative costs of the land in Bolivia. Including climate and geo-positioning makings help to determine the level and speed of economic progress, and subsequently the poverty, of a region. Bolivia presents forth a complicated but remarkable landscape that is divided between the lowlands and highlands. Joppa & Pfaff (2009), at the country-level, and Andam et al. (2010) and Sims (2010) have reported in their findings that protected areas are commonly found on steep and high ground. The slope of each municipality is measured with the average slope of each 1 ha patch therein. The data shows that protected areas do, in fact, have higher slopes but a lower elevation within the protected canton.

Forest Cover. Protected areas usually contain forested areas, therefore, Forest Cover of protected lands and Cantons alike were included. Due to the utility that forests represent as a source of income it is a significant variable to include in the study. This index is calculated from the percentage of forested areas for each canton.

Matching Estimators

Taking under consideration the wide array of methods available for the measurement of the effect of covariates of interest. For the covariate distribution it is implemented the nearest neighbor to each treated ATT unit to counteract the effect of each non-treated individual. These estimates are extracted from a relatively small sample size, hence the use of genetic matching is the method that utilizes a genetic algorithm that gives an optimal weight to each covariate while automatically locating the set of pairs that reduce the divergence between the distribution of potential confounders of control and treatment groups. This method produces an improved balance in the space of the metrics, rather than the incorporation of the Mahalanobis distance[2], the former is preferred; this relation is given by:

$$d(X_i, X_j) = \{(X_i - X_j)^T (S^{-1/2})^T W S^{-1/2} (X_i - X_j)\}^{1/2} \tag{10}$$

Where W is the weight positive definite $k * k$ matrix with all its elements equal to zero, except the main diagonal, S is the variance-covariance matrix of X and $S^{1/2}$ is the Cholesky decomposition (Sekhon, 2008) .

The latter technique incorporates the Mahalanobis distance between two column vectors with the addition of a weight matrix. When implementing genetic matching, the idea of the genetic algorithm is employed to find the previously mentioned improved balance. Other investigative works have implemented genetic matching with calipers to obtain increased levels of robustness in addition to other variations in the estimations that take into account regression based estimators (Canavire-Bacarreza & Hanauer, 2013) .

The former was applied to the dataset and also to acquire robust estimators. Replacement

[2]The most common method for matching is realized with the implementation of Mahalanobis distance, and it is determined by: $md(X_i, X_j) = \{(X_i - X_j)^T S^{-1} (X_i - X_j)\}^{1/2}$, where S is the variance matrix of X.

tools were utilized to reduce the bias when there were compatible individuals and risk is also added by increasing the variance within the matching process. (Imbens & Wooldridge (2009); Dehejia & Wahba (2002); Abadie & Imbens (2006)) were central in the development of genetic matching and essential in the calculation of heteroskedastic standard errors. To transcend the unbiased estimator to the estimation of the ATT, a matching estimator was specified between unprotected and protected areas in order to isolate the remaining variation between treatments.

Results

As seen in Table 3 the outcomes yielded by the analysis are arranged in accordance to find the ATT (Average Treatment Effect on the Treated)and for that purpose genetic matching is used as the central method of estimation, however in order to compare, a Welch two sample t-test was applied for the Naive impact estimate method was used to find the ATT, which measures the average differences between the treatment and control group in which it was tested for significant differences about their mean; Genetic matching with an added caliper filtering method allows for the purge of the observations outside of one standard deviation about the mean (1 standard deviation). The use of a caliper filter allows us to implement a narrow scope on the data and an increased precision of the yielded estimators. The respective resulting balances are found on Tables 8, 9, and 10. In these results, a good balance between the covariate interests. In table 11 the Rosembam upper bounds are shown for genetic matching for each outcome variable and year in consideration. Table 1 shows the summary statistics for outcome variables and covariates of interest by year and protected and unprotected groups.

There is not evidence provided by the analysis that can demonstrate the significance of protected areas on the unemployment rate through genetic matching and genetic matching with calipers for 2001 and 2012 periods. A difference between the mentioned methods of estimation is encountered for unemployment outcome variable in 2001 of 0.245 for the protected group and 0.423 for the unprotected group and a difference of 0.128 for ATT. Differences for matching in 2012 are 0.036, 0.036 and 0.000 for protected, unprotected and ATT respectively.

Table 1: Summary statistics for outcome variables and covariates of interest

Variable	Description	Status	Mean	St. Dev.	Min	Median	Max
Poverty index 1992	Asset-based poverty index for 1992	Unprotected	-1.314	1.731	-4.081	-1.751	5.821
		Protected	-1.285	1.854	-3.701	-1.722	5.299
Poverty index 2001	Asset-based poverty index for 2001	Unprotected	-0.230	1.753	-3.621	-0.579	6.227
		Protected	0.119	2.049	-3.053	-0.224	5.899
Poverty index 2012	Asset-based poverty index for 2012	Unprotected	1.404	1.746	-3.010	1.265	6.585
		Protected	1.840	1.966	-3.146	1.621	5.948
Unemployment 1992	Unemployment rate for 1992	Unprotected	0.917	2.066	0.000	0.346	45.455
		Protected	0.667	0.881	0.000	0.304	4.017
Unemployment 2001	Unemployment rate for 2001	Unprotected	2.416	3.065	0.000	1.520	27.586
		Protected	2.685	4.140	0.000	1.523	29.412
Unemployment 2012	Unemployment rate for 2012	Unprotected	0.696	1.002	0.000	0.483	11.765
		Protected	0.724	0.934	0.000	0.474	5.263
Informality 1992	% occupied on the informal sector for 1992	Unprotected	66.045	18.240	0.000	69.444	98.010
		Protected	63.619	18.421	16.143	67.067	95.238
Informality 2001	% occupied on the informal sector for 2001	Unprotected	70.433	15.004	12.658	73.684	100.000
		Protected	64.765	16.746	0.000	66.415	100.000
Informality 2012	% occupied on the informal sector for 2012	Unprotected	63.956	14.685	0.000	66.667	96.970
		Protected	61.083	13.640	28.678	62.613	89.744
Agricultural 1992	% Workers dedicated to agriculture, hunting, fishing, forestry, mining and quarrying for 1992	Unprotected	83.005	17.287	2.660	88.690	100.000
		Protected	81.927	19.943	3.943	89.918	100.000
Agricultural 2001	% Workers dedicated to agriculture, hunting, fishing, forestry, mining and quarrying for 1992	Unprotected	72.006	17.856	2.968	76.495	100.000
		Protected	66.429	21.643	3.887	73.718	100.000
Agricultural 2012	% Workers dedicated to agriculture, hunting, fishing, forestry, mining and quarrying for 1992	Unprotected	69.733	18.647	3.217	73.880	100.000
		Protected	67.483	22.849	3.644	73.799	97.826
Percent Forest Cover 1991	% of canton under forest cover in 1991	Unprotected	0.152	0.291	0.000	0.000	0.999
		Protected	0.396	0.376	0.000	0.338	0.977
Distance to Major City (km)	Avg. dist. to a city from each 1ha parcel within each canton	Unprotected	99.668	64.784	2.570	88.624	565.712
		Protected	139.017	118.168	5.382	117.895	589.060
Average Elevation (m)	Average elevation of each 1ha parcel within each canton	Unprotected	2,958.293	1,292.522	128.402	3,542.900	4,840.510
		Protected	2,295.522	1,454.963	136.618	2,444.215	4,838.710
Average Slope (pct)	Average slope of each 1ha parcel within each canton	Unprotected	20.843	15.064	0.797	19.413	66.924
		Protected	27.093	15.788	1.302	29.639	60.455
Roadless Volume (km)	Sum of the product of area and dist. to road, each 1ha parcel	Unprotected	22.463E+07	12.443E+08	1.258E+04	13.750E+06	19.200E+09
		Protected	1.172E+09	5.652E+09	26,864.1	43.784E+06	54.100E+09

Notes: Sample includes 106 protected cantons and 1108 unprotected cantos. 39 'marginally' protected cantons are removed from the sample (Protected unit between 0.01 and 0.1 percent)

The results are similar for both methods of estimation and it means that there is robust estimation of the genetic matching given the small variances between either methods of estimation, however not only protected areas do not significantly affect unemployment but in table 11 it is possible to observe that for both years of the analysis, there are other non observed factors that affect unemployment although it is an expected result because of the non-significance of protected areas on unemployment.

Table 2: Results from first stage matching for poverty samples

	Poverty index 2001			*Poverty index 2012*		
	Protected Y(1)	Unprotected Y(0)	ATT 2001	Protected Y(1)	Unprotected Y(0)	ATT 2012
Naïve diff. in means[a]	0.119	-0.204	0.323	1.840	1.437	0.403**
	[106]	[1146]	{0.119}	[106]	[1146]	{0.044}
Naïve diff. in means	0.119	-0.230	0.349	1.840	1.404	0.436**
	[106]	[1107]	{0.092}	[106]	[1107]	{0.030}
Genetic Matching	0.119	-0.270	0.389***	1.840	1.355	0.485***
	[106]	[106]	(0.131)	[106]	[106]	(0.189)

[a] Without excluding marginally protected units.

[Number of observations].

(Abadie and Imbens heteroskedasticity robust standard errors).

{P-value}.

***,** represent significance at the 1% and 5% level, respectively.

Labor informality brought forth contrasting results as there is significant evidence of positive impact between the establishment of protected areas on informality. Protected areas represent a reduction in the informality within the labor force across the nation in 4.535% and 4.523% in 2001 and 2012 respectively. The effect of protected areas on labor informality levels is consistent in time. All methods applied for the estimation assert a high level of significance linking informality and protected areas within the examined cantons.

There was larger observed difference in labor informality than in unemployment with respect to the two methods of estimation, genetic matching and genetic matching with calipers for this outcome variable in 2001 the differences are 1.492, 2.023 and 0.531 for protected, unprotected and ATT respectively. For 2012 the differences are 0.716, 1.355 and 0.639 for protected, unprotected and ATT respectively.

Table 3: Results from labor market

Analysis	Algorithm	*2001* Protected Y(1)	Unprotected Y(0)	ATT 2001	*2012* Protected Y(1)	Unprotected Y(0)	ATT 2012
Unemployment	Naïve impact estimate[a]	2.685	2.430	0.255	0.724	0.694	0.03
		[106]	[1146]	{0.538}	[106]	[1146]	{0.756}
	Naïve impact estimate	2.685	2.416	0.269	0.724	0.696	0.028
		[106]	[1107]	{0.516}	[106]	[1107]	{0.766}
	Genetic Matching	2.685	2.568	0.117	0.724	0.598	0.126
		[106]	[106]	(0.511)	[106]	[106]	(0.102)
	Genetic Matching with calipers	2.440	2.145	0.245	0.760	0.634	0.126
		[98]	[98]	(0.347)	[98]	[98]	(0.098)
Informality	Naïve impact estimate[a]	64.764	70.169	-5.405***	61.083	63.774	-2.691*
		[106]	[1146]	{0.002}	[106]	[1146]	{0.055}
	Naïve impact estimate	64.764	70.433	-5.669***	61.083	63.955	-2.872**
		[106]	[1107]	{0.001}	[106]	[1107]	{0.041}
	Genetic Matching	64.764	69.299	-4.535***	61.083	65.606	-4.523**
		[106]	[106]	(1.663)	[106]	[106]	(1.907)
	Genetic matching with calipers	66.256	71.322	-5.066***	61.799	66.961	-5.162***
		[95]	[95]	(1.776)	[95]	[95]	(1.993)
Agricultural	Naïve impact estimate[a]	66.429	71.833	-5.404**	67.483	69.504	-2.021
		[106]	[1146]	{0.014}	[106]	[1146]	{0.378}
	Naïve impact estimate	66.429	72.005	-5.576**	67.483	69.732	-2.249
		[106]	[1107]	{0.011}	[106]	[1107]	{0.327}
	Genetic Matching	66.429	73.178	-6.749**	67.483	72.127	-4.664**
		[106]	[106]	(2.549)	[106]	[106]	(2.352)
	genetic Matching with calipers	67.271	72.33	-5.059**	67.705	72.721	-5.016**
		[97]	[97]	(2.124)	[97]	[97]	(2.051)

[a] Without excluding marginally protected units.
[Number of observations].
(Abadie and Imbens heteroskedasticity robust standard errors).
{P-value}.
***,** represent significance at the 1% and 5% level, respectively.

In table 11, columns 3 and 4 are shown the Rosembaum upper bounds for genetic matching with informality as the outcome variable and the results obtained means that there are unobserved factors that affect the informality in a Γ higher than 1.1 and it is a good result, because estimates are still significant for Γ higher than 1.

In regards to the agricultural outcome variable a significant reductive impact is observed on the proportion of extractive economic activities as a result to the exposure of protected areas, extractive activities have been reduced in 6.749% in 2001 and 4.664% in 2012. There was a resulting difference of the protected of 0.842, a difference of the unprotected of 0.848, and a difference of ATT between methods of 1.69. Furthermore, there is robust estimation of the genetic matching giving the small variances between either methods of estimation. For the concerned outcome variable in 2012, we encountered a difference of the protected equal to 0.222, a difference of the unprotected of 0.594, and a difference of the ATT of 0.352. In Table 11, columns 6 and 7, it is observable that estimates remain significant until a Γ equal to 1.3 in 2001, so it is possible to conclude that extractive activities are impacted by protected areas while in 2012 there are unobservable factors that affected this reduction.

Discussion

Given the recent environmental movements to preserve land and biodiversity for the benefit of the commonwealth, it is imperative to estimate the effect that such decisions from the permanent geographical exclusion of human establishments have on the society as a whole. The critical measures of unemployment, informality, and extractive agricultural practices form a suitable reference to determine the socioeconomic impact that the public must endure in order to promote land preservation initiatives. The incorporation of different methods, as outlined in this work, are critical in asserting trustworthy conclusions. The purpose of this study is not to discredit any governmental bodies' decision to forgo the use of land, a critical economic factor central for growth and development, but rather offer insight into the externalities that such decisions have on society and their evolution over time. The necessity for protected areas to balance human concentrations is indisputable under any grounds; otherwise the mankind would assure catastrophic natural disasters as powerful forces that drive mother nature to correct excesses of human voracity in the use of its resources.

There is a wide evidence in previous studies, that protected areas affect poverty in a positive way, in table 2 the fist stage matching for poverty is shown, where evidence was found in a disaggregated by area (by canton instead of municipality). The estimation is robust for both outcome variables taken in consideration (PI 2001 and PI 2012) as the table

4 shows a Γ of 1.4 for 2001 and 2012. Balance results are shown in table 7, where it is evident that a good balance is achieved.

The mechanism in which poverty is affected by protected areas is trough labor market. There is evidence that informality decreases because of the establishment of protected areas, and changes in the workers' activities, from extractive activities, to others.

The mechanisms in place to fence-off land for protected areas is a profitable business venture for society as a whole, as well as a proven poverty reduction technique for local communities. There are tangible socioeconomic shifts such as reductions in labor informality, poverty, and unemployment along with significant increases in labor benefits, tax revenues, and tourism as a result of the goods and services that a natural reserve requires in its building and maintenance phases. The measurable impact that protected areas have on society is initially stronger than the next period as the effect marginally diminishes. Tourism could be the reason.

It is a significant channel to the reduction of poverty, as foreign funds rush to local communities, it suggests that further research should shed light on its poverty-fighting powers. If follow-up research asserts that tourism is an effective driving force to battle poverty, the government should consider instituting plans to propel the number of tourists currently visiting the surrounding protected areas. Government initiatives could be leveraged through the quantification and sale of carbon emission titles within the protected areas.

References

Abadie, A. & Imbens, G. W. (2006). Large sample properties of matching estimators for average treatment effects. *econometrica*, 74(1), 235–267.

Andam, K. S., Ferraro, P. J., & Hanauer, M. M. (2013). The effects of protected area systems on ecosystem restoration: a quasi-experimental design to estimate the impact of costa rica's protected area system on forest regrowth. *Conservation Letters*, 6(5), 317–323.

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. *Proceedings of the National Academy of Sciences*, 105(42), 16089–16094.

Andam, K. S., Ferraro, P. J., Sims, K. R., Healy, A., & Holland, M. B. (2010). Protected areas reduced poverty in costa rica and thailand. *Proceedings of the National Academy of Sciences*, 107(22), 9996–10001.

Bandyopadhyay, S. & Tembo, G. (2010). Household consumption and natural resource management around national parks in zambia. *Journal of Natural Resources Policy Research*, 2(1), 39–55.

Brockington, D., Igoe, J., & SCHMIDT-SOLTAU, K. (2006). Conservation, human rights, and poverty reduction. *Conservation Biology*, 20(1), 250–252.

Canavire-Bacarreza, G. & Hanauer, M. M. (2013). Estimating the impacts of bolivia's protected areas on poverty. *World Development*, 41, 265–285.

Dehejia, R. H. & Wahba, S. (2002). Propensity score-matching methods for nonexperimental causal studies. *Review of Economics and statistics*, 84(1), 151–161.

Ferraro, P. J. & Hanauer, M. M. (2011). Protecting ecosystems and alleviating poverty with parks and reserves:'win-win"or tradeoffs? *Environmental and Resource Economics*, 48(2), 269–286.

Ferraro, P. J., Hanauer, M. M., Miteva, D. A., Canavire-Bacarreza, G. J., Pattanayak, S. K., & Sims, K. R. (2013). More strictly protected areas are not necessarily more protective: evidence from bolivia, costa rica, indonesia, and thailand. *Environmental Research Letters*, 8(2), 025011.

Ferraro, P. J., Hanauer, M. M., & Sims, K. R. (2011). Conditions associated with protected area success in conservation and poverty reduction. *Proceedings of the National Academy of Sciences*, 108(34), 13913–13918.

Imbens, G. & Wooldridge, J. (2009). New developments in econometrics. *Lecture Notes, CEMMAP, UCL2009b.*

Joppa, L. N. & Pfaff, A. (2009). High and far: biases in the location of protected areas. *PloS one*, 4(12), e8273.

Sekhon, J. S. (2008). Multivariate and propensity score matching software with automated balance optimization: the matching package for r. *Journal of Statistical Software, Forthcoming.*

Sims, K. R. (2010). Conservation and development: Evidence from thai protected areas. *Journal of Environmental Economics and Management*, 60(2), 94–114.

Wade, T., Sommer, S., et al. (2006). *A to Z GIS, An illustrated dictionary of geographic information systems.* Esri Press.

Appendix

Table 4: Rosembaum upper bound Genetic Matching PI 2001 and PI 2012

	PI 2001	*PI 2012*
Γ	Upper bound P-value	Upper bound P-value
1.0	0.0016	0.0017
1.1	0.0059	0.0060
1.2	0.0162	0.0166
1.3	0.0366	0.0373
1.4	0.0703	0.0716

Γ represents the Gamma at which estimates are still significant at 10% level according to sensitivity to unobserved heterogeneity test.

Table 5: Eigen Vector Values of Variables used to create poverty Index

Pooled variable	Eigen V
Adult man in population[a]	0.052
Households without bathroom[a]	-0.307
Households that use fuelwood for cooking[a]	-0.338
Households with dirt floors[a]	-0.365
Low-quality houses[a]	-0.349
Households without electricity[a]	-0.509
Illiterate population[a]	-0.179
Population employed with salary[a]	0.167
Average persons per bedroom	-0.006
Households without access to public water[a]	-0.438
Households without sewer or septic[a]	-0.149
Average years of education	0.028

Notes: Census from 1992, 2001 and 2012 are pooled to measure average influence of assets across time.
[a] Indicates that variable is measured as a percentage.

Table 6: Mean Asset Values by Deciles of 1992 Poverty Index

Variable	Poorer			Deciles of PI 1992				Richer		
	1	2	3	4	5	6	7	8	9	10
Adult man in population	56.91	56.20	56.48	56.61	56.20	56.67	58.27	56.82	57.20	62.11
Without bathroom	98.61	95.19	91.36	91.03	84.30	82.93	80.15	77.51	64.40	23.72
Use fuelwood for cooking	96.80	92.92	88.86	86.50	76.81	76.16	71.85	70.89	52.59	5.89
With dirt floors	96.74	94.70	92.25	88.08	85.59	80.13	80.69	73.16	51.10	8.76
Low-quality houses	79.84	68.43	61.59	56.50	57.40	49.95	42.25	39.14	22.86	4.60
Without electricity	99.74	98.83	99.03	97.91	96.34	94.90	90.71	79.96	48.89	8.70
Illiterate population	57.77	43.91	36.34	34.05	33.70	30.14	30.51	27.07	21.01	12.10
Population employed with salary	3.85	8.57	10.35	11.27	15.96	14.61	18.25	21.61	36.92	69.51
Average persons per bedroom	3.83	3.64	3.50	3.40	3.55	3.34	3.24	3.32	3.30	2.92
No access to public water	97.66	95.37	93.10	90.05	88.41	85.27	78.92	69.98	47.10	10.40
Without sewer or septic	99.88	99.52	99.54	98.57	98.48	98.22	96.67	95.85	85.85	25.24
Average years of education	1.93	2.73	3.27	3.35	3.51	3.73	3.83	4.15	5.06	8.92

Table 7: Balance for genetic matching Poverty Index

Covariate	Status	Mean Prot.	Mean Unprot.	Diff.in means	Norm. diff[a]	Mean eQQ diff.[b]	% Improve mean diff.
Poverty index 1992	Unmatched	−1.285	−1.314	0.029	0.011	0.127	0.000
	Matched	−1.285	−1.298	0.013	0.005	0.084	0.535
% Forest 1991	Unmatched	0.396	0.152	0.244***	0.513	0.242	0.000
	Matched	0.396	0.404	−0.008	0.015	0.016	0.966
Distance to major city (km)	Unmatched	139.017	99.668	39.349***	0.292	41.318	0.000
	Matched	139.017	129.332	9.685	0.065	20.954	0.754
Average elevation	Unmatched	2295.522	2958.293	−662.772***	0.341	700.306	0.000
	Matched	2295.522	2284.619	10.903	0.005	104.620	0.984
Average slope	Unmatched	27.093	20.843	6.250***	0.286	6.247	0.000
	Matched	27.093	26.777	0.315	0.014	1.346	0.950
Roadless	Unmatched	1.17E+09	2.24E+08	9.47E+08*	0.164	8.61E+08	0.000
Volume 1992(km)	Matched	1.17E+09	8.22E+08	3.50E+08	0.054	4.62E+08	0.631

None of the post-match differences are significant.
[a]Normalized difference in means is the difference in means divided by the square root of the sum of the squared standard deviations of the treated and untreated covariate samples.
[b] Mean eQQ difference is the mean of the raw differences in the empirical quantile-quantile plots.
***,** rrepresent significant differences at the 1% and 5% level, respectively.

Table 8: Balance for genetic matching Unemployment rate

Covariate	Status	Mean Prot.	Mean Unprot.	Diff.in means	Norm. diff[a]	Mean eQQ diff.[b]	% Improve mean diff.
Unemployment rate 1992	Unmatched	0.667	0.917	-0.250^{**}	0.111	0.582	0.000
	Matched	0.667	0.681	-0.014	0.011	0.072	0.943
Poverty index 1992	Unmatched	-1.285	-1.314	0.029	0.011	0.127	0.000
	Matched	-1.285	-1.303	0.018	0.007	0.106	0.383
% Forest 1991	Unmatched	0.396	0.152	0.244^{***}	0.513	0.242	0.000
	Matched	0.396	0.403	-0.007	0.012	0.018	0.973
Distance to major city (km)	Unmatched	139.017	99.668	39.349^{***}	0.292	41.318	0.000
	Matched	139.017	128.205	10.812	0.073	26.162	0.725
Average elevation	Unmatched	2295.522	2958.293	-662.772^{***}	0.341	700.306	0.000
	Matched	2295.522	2190.156	105.365	0.053	123.290	0.841
Average slope	Unmatched	27.093	20.843	6.250^{***}	0.286	6.247	0.000
	Matched	27.093	25.972	1.121	0.050	2.465	0.820
Roadless	Unmatched	1.17E+09	2.24E+08	9.47E+08*	0.164	8.61E+08	0.000
Volume 1992 (km)	Matched	1.17E+09	8.55E+08	3.17E+08	0.049	4.83E+08	0.665

None of the post-match differences are significant.
[a]Normalized difference in means is the difference in means divided by the square root of the sum of the squared standard deviations of the treated and untreated covariate samples.
[b] Mean eQQ difference is the mean of the raw differences in the empirical quantile-quantile plots.
***,** represent significant differences at the 1% and 5% level, respectively.

Table 9: Balance for genetic matching Informality

Covariate	Status	Mean Prot.	Mean Unprot.	Diff.in means	Norm. diff[a]	Mean eQQ diff.[b]	% Improve mean diff.
Informality 1992	Unmatched	63.619	66.045	-2.426	0.094	3.287	0.000
	Matched	63.619	64.266	-0.647	0.025	1.732	0.733
Poverty index 1992	Unmatched	-1.285	-1.314	0.029	0.011	0.127	0.000
	Matched	-1.285	-1.270	-0.015	0.006	0.128	0.462
% Forest 1991	Unmatched	0.396	0.152	0.244^{***}	0.513	0.242	0.000
	Matched	0.396	0.392	0.004	0.007	0.015	0.984
Distance to major city (km)	Unmatched	139.017	99.668	39.349^{***}	0.292	41.318	0.000
	Matched	139.017	126.565	12.452	0.085	22.336	0.684
Average elevation	Unmatched	2295.522	2958.293	-662.772^{***}	0.341	700.306	0.000
	Matched	2295.522	2244.154	51.367	0.025	89.969	0.922
Average slope	Unmatched	27.093	20.843	6.250^{***}	0.286	6.247	0.000
	Matched	27.093	27.407	-0.314	0.014	0.973	0.950
Roadless	Unmatched	1.17E+09	2.25E+08	9.47E+08*	0.164	8.61E+08	0.000
Volume 1992 (km)	Matched	1.17E+09	7.58E+08	4.14E+08	0.065	4.68E+08	0.563

None of the post-match differences are significant.
[a]Normalized difference in means is the difference in means divided by the square root of the sum of the squared standard deviations of the treated and untreated covariate samples.
[b] Mean eQQ difference is the mean of the raw differences in the empirical quantile-quantile plots.
***,** represent significant differences at the 1% and 5% level, respectively.

Table 10: Balance for genetic matching % Agricultural

Covariate	Status	Mean Prot.	Mean Unprot.	Diff.in means	Norm. diff[a]	Mean eQQ diff.[b]	% Improve mean diff.
% agricultural 1992	Unmatched	81.927	83.005	−1.078	0.041	1.886	0.000
	Matched	81.927	83.420	−1.493	0.053	2.744	−0.385
Poverty index 1992	Unmatched	−1.285	−1.314	0.029	0.011	0.127	0.000
	Matched	−1.285	−1.297	0.012	0.005	0.094	0.565
% Forest 1991	Unmatched	0.396	0.152	0.244***	0.513	0.242	0.000
	Matched	0.396	0.405	−0.009	0.018	0.016	0.961
Distance to major city (km)	Unmatched	139.017	99.632	39.386***	0.292	41.343	0.000
	Matched	139.017	130.188	8.830	0.059	21.300	0.776
Average elevation	Unmatched	2295.522	2958.071	−662.549***	0.341	700.000	0.000
	Matched	2295.522	2249.135	46.387	0.023	99.667	0.930
Average slope	Unmatched	27.093	20.851	6.242***	0.286	6.242	0.000
	Matched	27.093	26.440	0.653	0.029	1.521	0.896
Roadless	Unmatched	1.17E+09	2.24E+08	9.47E+08*	0.164	8.61E+08	0.000
Volume 1992 (km)	Matched	1.17E+09	8.18E+08	3.54E+08	0.055	4.64E+08	0.627

None of the post-match differences are significant.
[a]Normalized difference in means is the difference in means divided by the square root of the sum of the squared standard deviations of the treated and untreated covariate samples.
[b] Mean eQQ difference is the mean of the raw differences in the empirical quantile-quantile plots.
***,** represent significant differences at the 1% and 5% level, respectively.

Table 11: Rosembaum upper bounds for Genetic matching

	Upper bound P- Value					
Γ	*Unempl. 2001*	*Unempl. 2012*	*Inform. 2001*	*Inform. 2012*	*% Agri.2001*	*%agri. 201*
1.0	0.652	0.118	0.028	0.026	0.005	0.123
1.1	0.792	0.220	0.068	0.065	0.014	0.232
1.2	0.885	0.346	0.134	0.129	0.035	0.364
1.3	0.941	0.480	0.224	0.217	0.071	0.503

Γ represents the Gamma at which estimates (Unemployment, Informality, %Agriculture) are still significant at 10%
sensitivity to unobserved heterogeneity test.

Table 12: Regression Results for Primary Specifications

	Dependent variable:		
	Protected		
	(1)	(2)	(3)
Unemployment rate 1992	-0.070		
	(0.057)		
Informality 1992		-0.001	
		(0.004)	
% agricultural 1992			-0.002
			(0.005)
Poverty index 1992	0.063^{*}	0.034	0.030
	(0.036)	(0.038)	(0.045)
% Forest 1991	1.043^{***}	1.029^{***}	1.036^{***}
	(0.269)	(0.268)	(0.267)
Distance to major city	0.002^{***}	0.002^{***}	0.002^{***}
	(0.001)	(0.001)	(0.001)
Average elevation	0.0001	0.0001	0.0001
	(0.0001)	(0.0001)	(0.0001)
Average slope	0.018^{***}	0.019^{***}	0.019^{***}
	(0.004)	(0.004)	(0.004)
roadless volume	0.000^{**}	0.000^{**}	0.000^{**}
	(0.000)	(0.000)	(0.000)
Constant	-2.398^{***}	-2.413^{***}	-2.339^{***}
	(0.290)	(0.349)	(0.480)
Observations	1,214	1,214	1,213
Log Likelihood	-314.877	-315.652	-315.601
Akaike Inf. Crit.	645.753	647.304	647.203
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 13: Branches of Economic Activity

	Years	
Activities	2001	2012
1	Agriculture, livestock, hunting, forestry	Agriculture, livestock, hunting, fishing, forestry
2	Fishing	Mining operations
3	Mining operations	Manufacturing Industry
4	Manufacturing Industry	Electricity, natural gas, vapor, and air conditioning
5	Electricity, natural gas, and hydroactivity	Hydroactivity and Residual water discharge, decontamination, and waste water treatment
6	Construction	Construction
7	Commerce	Wholesale and retail sales; automobile, motorcycle, and bike repairs
8	Hotels and restaurants	Transportation and storage
9	Transportation, storage, communications	Housing accommodation (hotels) and restaurants
10	Financial intermediaries	Information and communications
11	Realestate, renting, and coorporate services	Financial intermediaries and insurance services
12	Public administration	Real-estate
13	Education	Professional and technical services
14	Health and social services	Administrative and supporting services
15	Community, social, and personal services	Public administration, defense, and Mandatory Affiliation Social Security Plans
16	Household and Domestic Services	Educational Services
17	Foreign Territorial Organizational Services	Health and social services
18	-	Artistic, entertaining, and recreational activities
19	-	Other Services
20	-	Private household employment activities, non-differentiated household activities as producers of goods and services for their own benefit
21	-	Foreign Territorial Organizational Services