Environment and Development Economics 20: 283–310 © Cambridge University Press 2014
doi:10.1017/S1355770X14000461

Protected areas and economic welfare: an impact evaluation of national parks on local workers' wages in Costa Rica

JUAN ROBALINO
Department of Economics, Universidad de Costa Rica and Environment for Development Initiative, CATIE, 7170, CATIE, Turrialba, Costa Rica.
Tel: 2558-2522.
Email: robalino@catie.ac.cr

LAURA VILLALOBOS
Department of Economics, University of Gothenburg, Sweden.
Email: Laura.Villalobos@economics.gu.se

Submitted 28 February 2012; revised 4 July 2013, 19 April 2014; accepted 29 April 2014; first published online 19 June 2014

ABSTRACT. The number of protected areas around the world has significantly increased. However, the effects of this policy on the wellbeing of local households are still under debate. Using pre-treatment characteristics and household surveys with highly disaggregated geographic reference, we explore how national parks affect the wages of local workers in Costa Rica. We use matching techniques to control for the endogenous location of parks. We find that parks' effects on wages are, on average, positive and significant, but the magnitudes vary. Wages close to parks are higher for local workers living near tourist entrances. However, there is no robust evidence of positive effects for those close to parks but far away from tourist entrances. With our individual-level data, we also show that the positive effects on local households might not be as large as suggested by previous studies that use aggregated level data containing both local and immigrant households.

1. Introduction

In the last few decades, the number of protected areas around the world has significantly increased. Since 1990, the world's protected areas have increased by 58 per cent in number and 48 per cent in area (UN, 2012). Moreover, initiatives to expand the amount of protected land in developing countries are under way (e.g., REDD, the United Nations Collaborative Program on Reducing Emissions from Deforestation and Forest Degradation in Developing Countries). However, the debate continues over how these efforts affect local communities. Sorting through the effects of protected areas on local communities is important to determine whether they are affected positively or negatively and whether it is feasible to promote policies that contribute to both poverty reduction and conservation.

It has been argued that national parks may not have positive effects in nearby communities. Land-use restriction can lead to loss of employment, social differentiation, inequality and uncertainty over property rights (Fortin and Gagnon, 1999; Pfeffer *et al.*, 2001; List *et al.*, 2006; Robalino, 2007). A considerable amount of research has shown that welfare effects of parks are insignificant (Duffy-Deno, 1998; Lewis *et al.*, 2002, 2003) and, globally, there is no evidence that protected areas attract or repel human settlement (Joppa *et al.*, 2009).

However, it has also been argued that national parks could have positive effects on local communities. Recent rigorous empirical studies that use impact evaluation methods have found evidence that parks benefit local communities. They show that protected areas can increase household income (Mullan *et al.*, 2009) and alleviate poverty (Andam *et al.*, 2010; Sims, 2010; Canavire-Bacarreza and Hanauer, 2013).

Our research contributes to this discussion in two dimensions. First, we show conditions in which the effects on local welfare can be positive or insignificant for different areas of a park. Using the spatial reference of the observations, we identified workers close to park entrances, close to parks but far from entrances, and far away from the parks. This, in turn, allows us to look at the effects on wages in places where most tourism activities take place and compare these effects with areas that are close to parks but far from tourism.

Second, the analysis at the individual level allows us to obtain more precise and detailed conclusions using individual characteristics. One of the challenges in previous studies has been dealing with the fact that, within aggregated observations, the effects of protected areas on locals and immigrants cannot be separated. This implies the possibility that all the benefits might be captured by wealthier households that immigrated due to the park. However, we can distinguish workers based on the location where they were born and estimate the effects separately.

One of our empirical challenges was the fact that parks (and park entrances) are endogenously located (Pfaff *et al.*, 2009). This implies that the characteristics of people living close to parks and close to entrances can differ significantly from the rest of the rural population. To address this issue, we used a large data set with workers' and geographic characteristics and with counties' pre-treatment characteristics to control for different potential sources of bias.

Using standard regression analysis and different matching strategies, we found that, on average, the effect on wages of being close to parks is positive and significant for local workers. The estimated effects range from 7.59 to 10.29 per cent. Moreover, we also found that the effects vary according to proximity to the entrance of the park. The estimated effects of living close to the entrance of a park range from 7.14 to 10.37 per cent. However, there was no robust evidence of positive or negative effects for workers living far from the entrance. When using matching techniques, the estimated impacts far from the entrance were insignificant or marginally significant.

Previous research that analyzed aggregated data included both local and immigrant households in their analyses. Because we are able to separate them, we show the differences between considering only local workers and considering both local and immigrant workers. As we mentioned, the estimated benefits in wages from living close to the park range from 7.59 to 10.29 per cent for local workers but, when we include immigrants, the estimated effects range from 14.47 to 16.34 per cent. The differences in the estimated effects were statistically significant. Similarly, when we compare the estimated benefits of living close to the entrance of a park, we find that they range from 7.14 to 10.37 per cent for local workers, but, if we include immigrants, the estimated effects range from 16.01 to 19.55 per cent. These differences are also statistically significant. Using all workers in the aggregated data to estimate the effects of protected areas might overestimate the benefits that local households receive.

We then analyzed park effects by groups. We found that, while both females and males received better wages close to park entrances, the premium for females was significantly larger. These differences are larger than 9 per cent and are statistically significant in three out of four matching strategies. Also, the estimated effect of being close to the park entrance on wages is greater than 17 per cent and is statistically significant in areas that had low population density before the establishment of the protected area. As low population density is associated with low development levels, these results might suggest that parks have important effects in less developed areas.

We conclude that there is no evidence of negative effects of parks on wages of local workers, and that the size of the positive effects depends on geographical and individual characteristics. We use regression analysis and different matching techniques. We test different definitions of being close to parks. Results robustly show no evidence of negative effects of parks on wages. We also conducted a placebo analysis using the matched control group as if they were treated (placebo group) and found no significant effect. These placebo test results show that there is no evidence that suggests that our findings are a consequence of an unobservable factor.

This paper is organized as follows. In section 2 we discuss the Costa Rican context in which this study takes place. In section 3 we refer to data and how they were obtained. We discuss our empirical strategy in section 4. In section 5 we present the results. Conclusions are presented in section 6.

2. Background

2.1. Protected areas and labor market in Costa Rica

Costa Rica is a relatively small country of 51.100 km^2 and around 4.5 million people, 41 per cent of whom live in rural areas. This Latin American country has a long tradition of conserving its natural resources. Nearly 26 per cent of its land and 17 per cent of its coastal waters are under conservation regimes (SINAC, 2006). Half of the land that is protected and almost all of the marine protected areas are designated national parks. This is one of the strictest protection policies according to the International Union for the Conservation of Nature (IUCN) classification (IUCN, 1994).

Currently, Costa Rica has 28 national parks distributed all around the country. The first national parks were established in 1955, but most were created in the 1970s. The main objective of the Costa Rican national parks is to preserve natural resources *in situ*; as a result, human settlement and agricultural development are not allowed within a park's borders (SINAC, 2006).

These restrictions on agricultural activities could have negative effects on the rural economy. This is especially relevant in a country such as Costa Rica, where agriculture is a key economic activity. In 2007 agricultural production was 7 per cent of gross domestic product (GDP) and employed 13 per cent of the labor force. These numbers are even more relevant in rural areas where most of the protected areas are implemented. Land use restrictions could certainly reduce the local demand for agricultural labor, which in turn can reduce local wages.

However, tourism is also a highly dynamic economic activity in the country. Hotels and restaurants represented 4 per cent of 2007 GDP and employed 5 per cent of the labor force. Ecotourism, specifically related to protected areas, plays a central role within the tourism industry. In the last five years, tourists made more than 1 million visits to protected areas in Costa Rica (SINAC, 2006) and around 54 per cent of all foreign tourists in Costa Rica in 2007 visited a protected area. The implementation of parks, which leads to more ecotourism activities, could have affected positively the demand for workers and, therefore, wages.

There are other channels through which protected areas can have effects on wages. Higher levels of population around the parks as a consequence of tourism activities could lead to higher investments in infrastructure and human capital. More investment in roads, for instance, will reduce transport costs. This could also lead to more investment and increase the demand for labor and, therefore, local wages. Additionally, higher investment in education will increase the productivity of workers, which in turn could also increase wages.

Differentials in wages could lead to immigration to and emigration from protected areas, which could also, in turn, affect wages around the parks. In Costa Rica, there are no legal restrictions requiring firms to hire local workers or preventing workers from moving around the country. However, the fixed costs of migration and imperfect housing and credit markets are important factors limiting migration, which might reduce these secondary effects.

The mechanisms through which parks can affect local wages are numerous. Our goal is to estimate the net effect that national parks have on local workers' wages. Costa Rica is an excellent place to study these effects on local communities because it is a developing country where tourism and agriculture are central to rural development. Additionally, Costa Rica's vast and well-established conservation efforts offer a unique opportunity to evaluate their effects.

2.2. Impact evaluations of protected areas on local welfare

Methodological challenges to properly estimate the net effects of parks on local welfare have been discussed in Andam *et al.* (2010) and Pfaff and Robalino (2008). These authors argue that, in order to find the effect on people's welfare caused by the parks' presence, it is necessary to

https://doi.org/10.1017/S1355770X14000461 Published online by Cambridge University Press

adequately estimate what would have happened if the park had not been established. Andam *et al.* (2010) argue that most previous studies have lacked an adequate baseline or counterfactual. Recent studies, however, have attempted to address this issue.

Wittemyer *et al.* (2008) compare population growth in buffers of protected areas with population growth in rural areas. They conclude that, for different countries in Latin America and Africa, the population growth rate is higher in regions close to national parks. This is evidence, they argue, that parks actually have positive impacts on welfare. However, it is not clear if using rural areas as a unique criterion will lead to adequate control groups. There might be other differences that could explain these results. Moreover, Joppa *et al.* (2009) argue that Wittemyer *et al.* (2008) obtained these results by mixing two incompatible data sets.

Using a significantly larger set of variables to control for differences between protected and non-protected areas in Thailand, Sims (2010) found that protection increased average consumption and lowered poverty rates. This was an important contribution to the literature, as it was one of the first empirical analyses that addressed the issue of the endogenous location of protected areas when measuring impacts on social outcomes. The unit of analysis in that study is sub-districts. Therefore, it measured consumption and poverty at aggregate levels. Sims (2010) also acknowledged that improvements in social outcomes at aggregate levels could be explained by out-migration of poorer and/or in-migration of wealthier households. If this was the case, local people might not be better off.

Other researchers find similar results in other countries. Andam *et al.* (2010) find positive effects on poverty for Costa Rica and Thailand. They show that protected areas tend to be located in areas with higher poverty rates than average. After defining adequate control groups, they find evidence that protected areas have alleviated poverty. In contrast, Canavire-Bacarreza and Hanauer (2013) show that, in Bolivia, protected areas tend to be located in wealthier municipalities. However, after adequately defining a control group, they also find that protected areas reduced poverty.

Ferraro and Hanauer (2011) and Ferraro *et al.* (2011) show that the reductions in poverty that protected areas generate can differ according to distance to cities and slope of the land. They find that, in both Costa Rica and Thailand, poverty reduction is higher in areas with high slopes than in areas with low slopes. They also show that, at an intermediate distance to major cities, the reductions in poverty are higher.

These studies have also used aggregated units of analysis: sub-districts in the case of Thailand (Andam *et al.*, 2010; Ferraro *et al.*, 2011; Sims, 2010), census tracts in the case of Costa Rica (Andam *et al.*, 2010; Ferraro and Hanauer, 2011), and municipalities in the case of Bolivia (Canavire-Bacarreza and Hanauer, 2013). Therefore, one could argue that in-migration and out-migration could explain the positive effects that they find.

One important contribution of this paper is using individuals as the unit of analysis. By using individual data, we are able to separate those who were born in the community from those who migrated from somewhere else. This allows us to test whether local people actually benefit from protected areas.

Another important contribution is measuring new types of heterogeneous effects. We test whether the effects are evenly distributed or concentrated around tourist areas (park entrances). We also test whether areas with different population density could also determine the magnitude of the effects. We use population density as an indicator of local economic development before the parks were implemented. Finally, we also test whether woman and men are affected similarly. These are the first steps toward understanding how the protected areas affect local wellbeing.

3. Data

Our unit of analysis is economically active individuals who work in the private sector for more than 20 and less than 70 hours per week. Individual-level data were obtained from the household surveys (Encuestas de Hogares de Propósitos Múltiples [EHPM]),[1] conducted by the National Institute of Statistics and Census of Costa Rica (Instituto Nacional de Estadística y Censos [INEC]) from 2000 to 2007. Every year, the sample of households interviewed changes. The information about the location of households, however, was provided at the census tract level.[2] We focus only on workers from households located in rural census tracts, because that is where national parks are located. There are 36,557 individuals meeting these criteria in the survey.

EHPM surveys include information about workers' characteristics and wages. We used the logarithm of hourly real wages as a dependent variable (as in Heckman, 1974; Blanchflower *et al.*, 1996; Black and Strahan, 2001; von Wacher and Schmieder, 2009). Hourly real wages were obtained by deflating nominal monthly wages using the consumer price index calculated by Costa Rica's Central Bank (July 2006 = 100) and dividing by the number of hours worked per month. Demographic variables available from the EHPM are sex and age, which we call innate workers' characteristics.

Additionally, there is individual information about whether workers were born in the same county, or resided there two years earlier. Those individuals who live in the county where they were born are treated as local workers. Those individuals who were not born in the same county where they lived at the time of the survey are treated as immigrants.

In addition to the individual-level data, we included information at the district[3] level from the National Census in 1973 (also obtained from the

[1] This survey has been implemented since 1987. It is currently the official source of wage data in Costa Rica and has been used widely (e.g., Funkhouser, 1996, 1998). However, the survey might be subject to similar problems, such as bias due to self-reported information, as any other survey around the world.

[2] A census tract is a relatively small geographic unit that contains between 40 and 60 households.

[3] Districts are formed by census tracts.

INEC). In 1973 there were 407 districts. We use these variables to control for baseline conditions in that year. We included as controls population density, share of male population, share of population by decades of age, share of population born in the same county, share of population that lived in the same county for at least five years, and share of population within each economic activity category. If an individual is located in a district created after 1973, the information corresponding to the former district is assigned to that observation.[4]

Protected areas were mapped by the Geographic Information System Laboratory at the Instituto Tecnológico de Costa Rica. We rely on the location of census tracts, the smallest spatial reference available in the survey, to identify treated and untreated individuals. With the maps of protected areas and census tracts, we computed the minimum distance between each tract center[5] to each national park (linear distance). Then, we identified tracts close to a national park (treated) and far from a national park (untreated).

We had to drop observations with missing values in any of the individual-level observations and those that we could not match with pretreatment characteristics. After this, we are left with 33,975 observations. Then we eliminated all those individuals who lived in other protected areas that were not national parks so that the control group would not be affected by other policies. Therefore, we are left with 32,595 observations.

We also calculated the distance by road from each tract's center to each park's entrance. Park entrances are defined as the locations of a park ranger station, and therefore are places where tourists pay the fee to enter the park. When a park has more than one entrance, we include them all.

This allowed us to split the workers who live close to the park into two different groups: (1) those who live in a census tract within a 5 km buffer around the park and also within a 20 km distance by road to the nearest park entrance, which we call 'close to the entrance' (1,575 observations); and (2) those who live in a census tract within a 5 km buffer around the park, but located more than 20 km from the park entrance by road, which we call 'far from the entrance' (1,018 observations). In the untreated group, we placed individuals located in a tract more than 10 km from any national park, and we call them 'far from the park' (22,029 observations).[6]

We chose a 5 km buffer and 20 km distance by roads based on previous fieldwork analysis that showed that these are reasonable distances to start testing effects within the Costa Rican context (Villalobos, 2009). Nonetheless, to test robustness, we also use a 4 km and a 6 km ring, and 18 and 22 km distances from the park entrance.

[4] Information for four new districts was dropped because they were created out of several districts.

[5] To estimate distances by road, we used the center designated by the INEC, which corresponds to the most populated area in the tract.

[6] The difference between 32,595 and 24,622 (the sum of the different treatments and controls) are observations that lie in the 5–10 km distance from the national parks. We drop these observations to clearly differentiate between treated and untreated observations.

For both the group close to the entrance and the group far from the entrance, we restricted the sample to include only workers living close to a park that was established after 1973, because this is the date on which we measure the baseline. However, to improve the control group, we restricted that group to workers living far from any national park, irrespective of the establishment date.

We also used geographic variables at the census tract level. We calculated average slope, average precipitation, and average elevation per census tract using geographic information systems. We were also able to calculate distances from the census tract to San José,[7] and to the closest health and education center. The density of different types of roads was also calculated per census tract.

4. Empirical approach

Randomly located parks and randomly located entrances of the parks would eliminate many of the possible biases of estimating their effects. If this were the case, we would only need to compare the wages of workers close to parks (or close to the entrances) with the wages of workers who live far from parks. Worker characteristics would be equal in expectation and the only reason for difference in wages would be the effect of parks on the labor market.

However, policies are rarely applied randomly and national parks and land-conservation policies are no exception (Pfaff and Robalino, 2008; Pfaff *et al.*, 2009). For instance, parks might be located in areas with low opportunity costs. Workers in these areas might be less educated and less productive and, therefore, receive lower wages. This implies that differences in wages might not be due to parks but due to other characteristics of the areas where parks were implemented. These issues create sample selection bias (Heckman, 1979; Rosenbaum and Rubin, 1983; Dehejia and Wahba, 2001). This is what we find in our data.

In table 1, we compare the three groups of workers: in the first column, we present the characteristics of workers far from parks and compare them with workers close to parks (CP) in column 2, with workers close to the entrance of parks (CE) in column 3, and with workers close to the park but far from the entrance (FE) in column 4. We found significant differences for many of the characteristics, starting with wages. Workers living close to park entrances receive higher wages than workers living far from parks. Also, workers living far from the entrances have lower wages than workers far from parks. However, as discussed, the wages of these groups may be different, not only due to the effects of parks, but also due to differences in individual and geographic characteristics. For instance, there is, on average, more female participation in the labor force, fewer Costa Ricans, and higher immigration in areas close to the entrance of the parks than in rural

[7] We choose only San José as an important market because, in 1973, all other current relevant markets and population centers were undeveloped.

Table 1. *Comparison of living far from parks and close to parks (close to and far from entrances) on selected characteristics*

Variable	(1) *FP (controls)*	(2) *Difference (CP-FP)*	(3) *Difference (CE-FP)*	(4) *Difference (FE-FP)*
Number of observations	22,029	2593	1575	1018
log wage (CRC*** per hour)	6.37	0.11***	0.22***	−0.06***
Workers' innate characteristics				
Male participation (%)	81.95	−5.51***	−7.86***	−1.89
Age	32.55	−0.37	0.25	−1.33***
Costa Rican (%)	88.57	−4.18***	−3.23***	−5.66***
People living in the same place for at least 2 years (%)	95.03	−2.09***	−3.54***	0.15
Characteristics of the county in 1973				
Population density (people/km^2)	154.71	−57.34***	−3.32	−140.91***
People born in the same county (%)	61.95	3.19***	11.09***	−9.03***
People living in the same county for at least 5 years (%)	76.87	5.63***	12.14***	−4.44***
Male participation (%)	52.13	0.18***	−0.15**	0.68***
19 years old and younger population (%)	58.53	0.99***	−0.06	2.60***
Population between 20 and 59 years old (%)	36.51	−0.59***	−0.17**	−1.25***
Population older than 60 (%)	4.96	−0.39***	0.23***	−1.35***
Agriculture, hunting, forestry and fishing (%)	64.27	4.47***	−2.05***	14.56***
Mining and quarrying (%)	0.88	−0.70***	−0.71***	−0.67***
Manufacturing (%)	6.63	−1.62***	0.56***	−4.99***
Utilities (electricity, gas and water) (%)	0.77	−0.15***	0.14***	−0.59***
Construction (%)	4.95	0.20**	1.01***	−1.06***
Wholesale and retail trade and restaurants and hotels (%)	6.63	−1.02***	0.17*	−2.87***
Transport, storage and communication (%)	3.19	−1.38***	−1.18***	−1.68***
Financing, insurance, real estate and business services (%)	1.38	−0.39***	−0.61***	−0.05
Community, social and personal services (%)	11.29	0.59***	2.68***	−2.66***
Geographic characteristics				
Density of primary roads (km/km^2)	0.15	0.03***	0.09***	−0.06***

(*continued*)

https://doi.org/10.1017/S1355770X14000461 Published online by Cambridge University Press

Table 1. *Continued.*

Variable	(1) *FP (controls)*	(2) *Difference (CP-FP)*	(3) *Difference (CE-FP)*	(4) *Difference (FE-FP)*
Slope	8.87	2.00***	1.88***	2.17***
Precipitation (mm)	3127.11	395.29***	−166.73***	1264.84***
Distance to San José (km)	10.82	−0.03	−0.13***	0.13***

Notes: ***CRC, Costa Rican colones; CRC 500, US$1 (May 2010).

areas far away from the parks. Additionally, on average, there are more foreign and younger workers far from the entrances, than in rural areas away from parks.

We also compared the characteristics of the counties in 1973, to test whether pre-treatment characteristics were significantly different among the groups. We found no statistically significant difference between the population density of counties far from parks and those close to park entrances, while counties far from entrances were significantly less dense than counties far from parks. Also, in that year, there was more native population and immigration in counties close to entrances than far from parks. On the contrary, there was less native population and immigration in counties far from entrances compared to far from parks. Female participation in the labor force was also higher close to park entrances and lower far from entrances, compared with other rural areas far from parks.

The structure of economic activities was, to some extent, similar among groups in 1973 (baseline conditions). The primary sector that includes agriculture, hunting and fishing is the most important in all cases, with around a 60 per cent share. However, the share of the primary sector was significantly higher far from entrances. There are also geographic differences. Areas close to entrances have higher density of primary roads, lower precipitation levels, are located closer to San José, and have steeper slopes than rural areas far from parks. Areas close to parks but far from the entrances also have steeper slopes. However, these areas have lower density of primary roads, higher precipitation and are located farther from San José than other rural areas far from parks. This is evidence that there are not only important differences between areas located close to and far from parks, but also, within areas close to parks, between those that are close to and far from the entrances. This suggests that evaluating park effects close to and far from the entrances can be highly relevant.

4.1. Addressing the selection bias problem

We address the non-random establishment of parks and park entrances by using matching methods to pre-process the data (Ho *et al.*, 2007). In particular, we use propensity score matching, which is useful for estimating treatment effects in observational studies when the dimensionality of the observable characteristics is high (Rosenbaum and Rubin, 1983; Dehejia and Wahba, 2001). The goal is to find an adequate untreated control

https://doi.org/10.1017/S1355770X14000461 Published online by Cambridge University Press

group that is similar to the treated group in all relevant pre-treatment characteristics. Similarity is defined in terms of the propensity score, which is the conditional probability of assignment to a particular treatment, given a vector of observed covariates (Rosenbaum and Rubin, 1983).

The advantage of using propensity score matching is that it is possible to determine how well the treatment and control groups overlap. As shown by Rosenbaum and Rubin (1983), overlap in the propensity scores is sufficient to remove bias from observed covariates. Therefore, estimations are less sensitive to the choice of functional form in the model (Rosenbaum and Rubin, 1983; Dehejia and Wahba, 2001). Another advantage is that the variance of the estimate of the average treatment effect will be lower in matched samples, compared with random samples, because the distributions of the covariates in the treated and control groups are more similar in matched samples than in random samples (Rosenbaum and Rubin, 1983). A third advantage is that, unlike standard techniques, matching avoids extrapolation to portions of covariate space where there are no data.

However, as with all approaches, matching requires assumptions about unobserved factors for the correct identification of the effect. There must not be unobservable factors that affect the outcome and that are simultaneously correlated to the presence of treatment. Another issue with matching is that there can be a decrease in the number of observations because unmatched observations are dropped. We argue that the rich set of available data helped us minimize the possibility of unobservable bias and that the sample size is large enough to permit this loss of observations and these degrees of freedom.

When defining how observations are matched, there is a tradeoff between the quality of the matches (how similar they are) and the number of observations included (and hence the precision of the estimates). To account for both concerns, we use two different strategies: (1) Epanechnikov kernel, in which all observations are included but higher weights are given to the most similar matches; and (2) a combination of caliper matching and nearest neighbor matching, where we specify a maximum number of matches per treated observation and impose a tolerance level on the maximum propensity score distance. This guarantees that only the most similar observations are kept. For this, we tested different combinations of number of neighbors and caliper: up to four matches with a caliper of 0.01 and also 0.001, and then one-to-one matching with a caliper of 0.01. After selecting comparable groups, we use two different methods to compute differences. In the first method, we use a regression with weights based on the number of matched untreated observations from each of the matching strategies, following Hill *et al.* (2003). In the second method, we select the observations that appear as adequate controls but use them once, without weighting, even if they appear many times as adequate matches. Results are very similar between these methods. This second strategy allows us to test the differences between the parameters of different regression specifications.

In order to account for any remaining differences in covariates, we run a standard OLS regression after matching, using only the matched observations to finally obtain the effects. This approach of pre-processing the data using matching and then running a parametric procedure is more reliable than running the OLS regression alone (Ho *et al.*, 2007). However, we present both results for comparison purposes.

4.2. Likelihood of being treated (propensity scores)

We ran a probit regression in order to estimate the conditional probability of being assigned to each treatment group: being close to a park (CP), being close to the park and an entrance (CE), and being close to a park but far from the entrance (FE) (table 2). This was done for the entire sample and for only those who were born in the same locality (local workers). We included as explanatory variables those worker and geographic characteristics that are not affected by the treatment but that might determine it, as well as pre-treatment characteristics of the counties.

Specifically, we included the innate workers' characteristics: sex and age, and, if applicable, also the nationality and a dummy variable indicating whether the person has lived in the same place for at least two years. As geographic variables, we included density of primary roads, slope, precipitation and distance to the capital city (in log). Finally, as baseline conditions, we included the characteristics of the counties in 1973. These include population density, share of male population, share of population by age groups, share of population born in the same county, share of population that lived in the same county for at least 5 years, and share of population within each category of economic activity.

When we consider the whole sample, we find that worker's innate characteristics, county characteristics in 1973 and geographic characteristics are significant factors associated with the probability of being close to a national park and close to a national park's entrance. Geographic characteristics and 1973 characteristics are also significant factors associated with being close to a park but far from the entrance. However, workers' innate characteristics seem not to be significantly associated with being close to the park but far from the entrance. This is also evidence of the importance of separating the effects of being close to the park from the effects of proximity to the entrance.

When we consider only local workers, many county characteristics in 1973 and geographic characteristics are significant factors associated with the probability of being close to a national park, either close to or far from a national park's entrance. Separating the effects of being close to the park by proximity to the entrance is also important for locals.

4.3. Evidence of comparable groups

We used two strategies to examine whether we had comparable groups. First, we checked whether there was enough overlap between the treated and the control group before and after matching. Then, we verified whether matching was effective in obtaining similar samples by observing the balance in the confounder variables between the treated and the control groups before and after the matching.

To check for overlap, we plotted the histograms of the propensity scores of the treated and untreated groups before matching, and treated and matched groups after matching. We did this both for the 'close to entrances'

Table 2. *Likelihood of being treated (probit regression)*

	(1)	(2)	(3)	(4)	(5)	(6)
	Local and immigrant workers			*Local workers*		
Variable	*CP*	*CE*	*FE*	*CP*	*CE*	*FE*
Workers' innate characteristics						
Sex (male)	−0.2056***	−0.2312***	−0.0519	−0.2419***	−0.1925***	−0.3298**
Age	0.0118**	0.0218***	0.0068	0.0027	0.0066	−0.0126
Age^2	−0.0001**	−0.0003***	−0.0002	0.0000	−0.0000	0.0003
Costa Rican	−0.0020***	−0.0021***	0.0003			
People living in the same place for at least 2 years	−0.0028***	−0.0037***	0.0021			
Characteristics of the county in 1973						
Population density (people/km^2)	−0.0015***	−0.0001	−0.1523***	−0.0019***	−0.0000	−0.2280***
People born in the same county	0.0046**	−0.0029	−0.0731***	0.0004	0.0021	−0.0561***
People living in the same county for at least 5 years	0.0058***	0.0726***	0.0221***	0.0079***	0.0542***	0.0146*
Male participation	0.0305***	0.2263***	−0.8291***	−0.0288*	0.1513***	−1.1154***
19 years old and younger population	0.0409***	0.0310**	0.9681***	0.0469***	−0.0039	0.8386***
Population between 20 and 59 years old	0.0286	0.0514**	1.1481***	0.0492*	0.0051	0.9102***
Agriculture, hunting, forestry and fishing	−0.0548***	−0.0445***	−0.2422***	−0.0545***	−0.0232***	−0.1786***
Mining and quarrying	−0.6644***	−0.3292***		−0.6034***	−0.2913***	
Manufacturing	−0.0692***	−0.0506***	−0.1660***	−0.0595***	−0.0085	0.0564
Utilities (electricity, gas and water)	−0.1053***	−0.0563***	0.2204*	−0.0400**	0.0382**	−0.0413
Construction	0.0220***	−0.0107	0.5343***	0.0455***	0.0301***	0.3444***

(*continued*)

Table 2. *Continued*

	(1)	(2)	(3)	(4)	(5)	(6)
	Local and immigrant workers			*Local workers*		
Variable	*CP*	*CE*	*FE*	*CP*	*CE*	*FE*
Wholesale and retail trade and restaurants and hotels	−0.1133***	−0.0107	−0.9557***	−0.1733***	−0.0174	−0.6413***
Transport, storage and communication	−0.2387***	−0.2484***	−0.4512***	−0.2276***	−0.1892***	−0.3448***
Finance, insurance, real estate and business services	0.0888***	0.0452**	0.0207	0.0241	−0.0212	−0.1584***
Geographic characteristics						
Density of primary roads	−0.0312	0.1079***	0.2758**	0.4328***	0.4085***	0.2802*
Slope	0.0022	−0.0094***	0.1480***	−0.0020	−0.0112***	0.1126***
Precipitation	0.0009***	0.0009***	0.0003	0.0011***	0.0014***	−0.0003
Precipitation^2	−0.0000***	−0.0000***	0.0000***	−0.0000***	−0.0000***	0.0000***
log distance to San José	−0.2458***	−0.1555***	−0.6083***	−0.1781***	0.0375	−1.7770***
Constant	−0.5938	−18.6734***	−29.8824***	−0.1952	−15.4715***	11.5445
Observations	24,622	23,604	23,047	12,626	12,113	11,829
Prob > chi^2	0.000	0.000	0.000	0.000	0.000	0.000

Notes: $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Standard errors in brackets. Also controlled by year dummies.

Table 3. *Balances in characteristics before and after matching for local workers*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Close to park			*Close to entrance*			*Far from entrance*		
Variable	*Before matching*	*After matching: kernel*	*After matching: $n = 4$, $cal = 0.001$*	*Before matching*	*After matching: kernel*	*After matching: $n = 4$, $cal = 0.001$*	*Before matching*	*After matching: kernel*	*After matching: $n = 4$, $cal = 0.001$*
log wage	0.034**	0.113***	0.082***	0.092***	0.080***	0.073***	−0.057**	−0.035	0.051
Workers' characteristics									
Male participation (%)	−0.048***	0.059***	0.014	−0.047***	−0.020	0.007	−0.050***	0.202	0.092
Age	0.456	−0.972	−0.687	1.563***	0.382	0.708	−1.263**	2.525	−1.062
Characteristics of the county in 1973									
Population density (people/km^2)	−65.560***	−0.972	1.366	−32.530***	3.471	−13.047	−116.875***	1.894	2.274
People born in the same county (%)	0.260	2.355**	0.403	6.940***	6.724***	0.776	−10.118***	27.480***	1.332
People living in the same county for at least 5 years (%)	3.192***	0.070	0.137	8.448***	3.734***	0.679**	−4.975***	42.104***	5.175

Table 3. *Continued.*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Close to park			*Close to entrance*			*Far from entrance*		
Variable	*Before matching*	*After matching: kernel*	*After matching: n = 4, cal = 0.001*	*Before matching*	*After matching: kernel*	*After matching: n = 4, cal = 0.001*	*Before matching*	*After matching: kernel*	*After matching: n = 4, cal = 0.001*
Male participation (%)	0.528***	−0.334***	−0.392***	0.345***	−0.091	−0.164**	0.814***	1.184	0.454*
19 years old and younger population (%)	0.980***	−0.500***	0.201	0.251**	−0.100	0.153	2.113***	−1.345	0.142
Population between 20 and 59 years old (%)	−0.463***	0.311**	0.074	−0.219**	−0.138	−0.268	−0.842***	1.370	0.088
Population older than 60 (%)	−0.517***	0.190**	0.126*	−0.032	0.239***	0.115	−1.271***	−0.025	−0.230
Agriculture, hunting, forestry and fishing (%)	6.476***	−2.683**	−2.050**	1.421*	−2.059*	0.568	14.329***	−7.676	−2.011
Mining and quarrying (%)	−0.762***	−0.114	−0.033	−0.734***	−0.160	0.000	−0.804***	−0.251	−2.427**

Manufacturing (%)	−2.308***	0.501	0.489*	−0.801***	0.381	−0.053	−4.649***	0.928	−0.178
Utilities (electricity, gas and water) (%)	0.193***	−0.038	−0.190***	0.675***	0.794***	−0.209***	−0.554***	0.055	0.076
Construction (%)	0.045	0.731***	0.187	0.810***	1.070***	−0.222	−1.144***	1.530	0.304
Wholesale and retail trade and restaurants and hotels (%)	−1.351***	0.544***	0.473***	−0.369***	−0.286	−0.125	−2.876***	0.743	0.308
Transport, storage and commu-nication (%)	−1.306***	0.135	0.173**	−1.153***	−0.168	−0.029	−1.544***	0.895	0.189
Financing, insurance, real estate and business services (%)	−0.442***	0.124	0.069	−0.492***	−0.118	0.003	−0.365***	0.012	0.331
Community, social and personal services (%)	−0.545**	0.798**	0.883***	0.644**	0.546	0.068	−2.391***	3.765	3.408***

https://doi.org/10.1017/S1355770X14000461 Published online by Cambridge University Press

Table 3. *Continued*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Close to park			*Close to entrance*			*Far from entrance*		
Variable	*Before matching*	*After matching: kernel*	*After matching: n = 4, cal = 0.001*	*Before matching*	*After matching: kernel*	*After matching: n = 4, cal = 0.001*	*Before matching*	*After matching: kernel*	*After matching: n = 4, cal = 0.001*
Geographic characteristics									
Density of primary roads (km/km^2)	0.056***	0.024	0.070***	0.105***	−0.424***	−0.038	−0.020	0.023	−0.020
Slope	1.196***	1.738***	1.389***	0.463	1.969***	−0.002	2.334***	9.230**	0.471
Precipitation (mm)	625.353***	−22.832	−71.948*	183.160***	−178.709***	30.745	1312.348***	522.292	273.467
Distance to San José (km)	0.053**	−0.176***	−0.106***	0.060*	−0.163***	0.062	0.041	−0.587*	−0.050
Total unbalanced control variables at 5%	19	12	12	20	10	4	20	4	3

Notes: *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. Standard errors in brackets.

and 'far from entrances' analyses. For the 'close to entrances' analysis, the distributions of the treated and untreated groups are significantly different before matching. However, after matching, the distributions are more similar (figures are available in the online Appendix available at http://journals.cambridge.org/EDE). The difference between before and after matching is more striking when looking at the group far from entrance. There are some intervals where there are not enough matches and therefore we could not consider them in the analysis. For these intervals, we lack empirical evidence to properly estimate the treatment effects.

In the balance test, we found that, for most of the covariates, the differences between treated and controls are substantially reduced after matching (table 3). The matching is particularly effective when we disaggregate the analysis by distance to the entrance. This is consistent with the differences in the characteristics between the groups discussed previously. The number of covariates for which there are statistically significant differences is reduced from 20 before matching to four for the group close to the entrance and to three for the group far from the entrance. For the group far from the entrance, even when differences remain, matching improved similarity. Also, it can be seen that nearest neighbor matching with four matches and a small caliper (0.001) performs better in terms of balances with respect to kernel matching for the analysis close to the entrance. However, the matching strategies perform similarly for the treatment far from the entrance. These results show that the treated and the control groups are more comparable after matching than before.

5. Results

We use the log of hourly real wages as a dependent variable (as in Heckman, 1974; Blanchflower *et al.*, 1996; Black and Strahan, 2001, von Wacher and Schmieder, 2009). Therefore, the coefficient might be interpreted as the percentage change in the hourly wage caused by the treatment. The treatments we test are: (i) being close to the park; (ii) being close to the entrance of the park; and (iii) being close to the park but far from the entrance.

In table 4, we compare the effects estimated through different methodologies for local workers (those who were born in the same location). First, we estimated the naïve regression (Morgan and Winship, 2007), which in this case is a standard ordinary least squares (OLS) regression, controlling by year dummies. The results indicate that, on average, wages close to parks are 3.57 per cent higher than wages far from parks, but those differences are not evenly distributed around the park. Local workers who are close to parks' entrances earn 9.32 per cent more than local workers far from parks. Local workers who are far from entrances earn 5.4 per cent less than local workers far from parks. As discussed, these differences could be the result of differences in pre-treatment, geographic and innate individual characteristics, and/or could be caused by the effects of the treatment (the presence of parks).

Therefore, we control for those variables using standard OLS and four different matching estimators: kernel, one-to-four with a caliper of 0.001,

Table 4. *National parks' effects on local workers' wages*

Model		(1) *Close to park*	(2) *Close to entrance*	(3) *Far from entrance*	(4) *Difference (3) vs. (2)*
Naive regression (OLS)[a]	Effect	0.0357**	0.0932***	−0.0540**	0.1471***
	Standard error	[0.015]	[0.018]	[0.026]	
Standard regression (OLS)	Effect	0.0860***	0.1035***	0.0521*	0.0514*
	Standard error	[0.016]	[0.018]	[0.027]	
Standard regression with matching (kernel)	Effect (weights)	0.0957***	0.0846***	−0.0108	
	Standard error	[0.018]	[0.019]	[0.069]	
	Effect (no weights)[b]	0.0818***	0.1037***	0.0515*	0.0521
	Standard error	[0.016]	[0.019]	[0.030]	
Standard regression with matching ($n = 4$, cal = 0.001)	Effect (weights)	0.0816***	0.0721***	0.1782*	
	Standard error	[0.020]	[0.023]	[0.093]	
	Effect (no weights)	0.0759***	0.0749***	0.1219	−0.047
	Standard error	[0.019]	[0.022]	[0.093]	
Standard regression with matching ($n = 4$, cal = 0.01)	Effect (weights)	0.0881***	0.0714***	−0.0112	
	Standard error	[0.020]	[0.023]	[0.080]	
	Effect (no weights)	0.0789***	0.0797***	0.0128	0.0669
	Standard error	[0.019]	[0.021]	[0.066]	
Standard regression with matching ($n = 1$, cal = 0.01)	Effect (weights)	0.0915***	0.0768**	0.1097	
	Standard error	[0.027]	[0.031]	[0.094]	
	Effect (no weights)	0.1029***	0.0772***	0.0492	0.028
	Standard error	[0.025]	[0.028]	[0.092]	

Notes: *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. Standard errors in brackets.
[a] Using only year dummies as control variables. All other specifications include as control variables: workers' innate characteristics, geographic characteristics and characteristics in 1973.
[b] Regression on the restricted sample, without including weights for observations that appear more than once as adequate controls.

one-to-four with a caliper of 0.01, and one-to-one with a caliper of 0.01. After controlling for pre-treatment, geographic and workers' innate characteristics, the estimated wage effects of being close to parks are positive and significant. The estimated impact ranges from 7.59 to 10.29 per cent (column 1 in table 4). The estimated impacts of living close to the entrance range from 7.14 to 10.35 per cent (column 2 in table 4).

https://doi.org/10.1017/S1355770X14000461 Published online by Cambridge University Press

In contrast, the estimates of the effect of being far from the entrance shifted significantly after using the rest of the controls. They are now either insignificant or marginally significant. The standard OLS estimate is marginally significant and positive. However, six of the matching estimators show a non-significant effect and two of them are negative (column 3 in table 4). Therefore, we conclude that there seems to be no robust evidence of a positive or negative wage effect far away from the entrance. Consistent with these results, there is a lack of robust evidence when testing whether differences between those living close to and far from the entrance are statistically significant (column 4).

Our individual-level data allow testing of the effects on local people. This is an important contribution with respect to previous research that analyzed aggregated data that included both local and immigrant workers. In table 5, we show the differences between considering only local workers and both local and immigrant workers using matching estimators. The estimated effect on wages from living close to the park ranges from 7.59 to 10.29 per cent for local workers but, when we include immigrants, the estimated effects range from 14.47 to 16.34 per cent. We find that differences in estimates between these two groups are larger than 6 per cent and are always statistically significant.

Similarly, when we compare estimates of the benefits of living close to the entrance, we find that they range from 7.21 to 10.37 per cent for local workers but, if we include immigrants, the estimated effects range from 16.01 to 19.55 per cent. We also find that differences in estimates between these two groups are larger than 9 per cent and are always statistically significant. Using all workers in the aggregated data to estimate the effects of protected areas might overestimate the benefits that local workers perceive. However, we still find that the effects of protected areas on wages are positive and significant.

Finally, we find that the estimated effects for those living close to the park but far from the entrance are mostly insignificant and non-robust for both groups analyzed.

Taking advantage of our individual-level data, we also investigate other differences in the estimated effects. We analyze the difference in the premiums for living close to the entrance by sex and pre-treatment population density (see table 6). We found that, although both females and males receive better wages close to park entrances, the premium for local females (from 15.52 to 20.18 per cent) is larger than the premium for local males (from 6.47 to 7.23 per cent). For all matching strategies, the difference is always larger than 9 per cent. Additionally, this difference is statistically significant in three out of four matching methods.

Next, we estimate the effect according to the pre-treatment population density as a proxy for pre-treatment economic development of the county.[8] We focus this analysis on local workers close to the entrance. We use the median value of the sample to split the data between counties with low and

[8] We use population density as a proxy for pre-treatment development due to the lack of pre-treatment income and production data.

Table 5. *Results for local workers only and for local and immigrant workers, by matching strategy*

	(1) *Local workers (weights)*	(2) *Local and immigrant workers (weights)*	(3) *Local workers (no weights)*	(4) *Local and immigrant workers (no weights)*	(5) *Difference (4) vs. (3)*[a]
Close to parks					
Kernel	0.0957***	0.1447***	0.0818***	0.1536***	0.0718***
Standard error	[0.018]	[0.014]	[0.016]	[0.011]	
$n = 4$, caliper $= 0.001$	0.0816***	0.1549***	0.0759***	0.1548***	0.0789***
Standard error	[0.020]	[0.016]	[0.019]	[0.014]	
$n = 4$, caliper $= 0.01$	0.0881***	0.1551***	0.0789***	0.1579***	0.0790***
Standard error	[0.020]	[0.016]	[0.019]	[0.013]	
$n = 1$, caliper $= 0.01$	0.0915***	0.1597***	0.1029***	0.1634***	0.0605**
Standard error	[0.027]	[0.020]	[0.025]	[0.018]	
Close to entrances					
Kernel	0.0846***	0.1868***	0.1037***	0.1955***	0.0919***
Standard error	[0.019]	[0.016]	[0.019]	[0.014]	
$n = 4$, caliper $= 0.001$	0.0721***	0.1819***	0.0749***	0.1923***	0.1174***
Standard error	[0.023]	[0.018]	[0.022]	[0.017]	
$n = 4$, caliper $= 0.01$	0.0714***	0.1760***	0.0797***	0.1922***	0.1125***
Standard error	[0.023]	[0.018]	[0.021]	[0.017]	
$n = 1$, caliper $= 0.01$	0.0768**	0.1601***	0.0772***	0.1733***	0.0961**
Standard error	[0.031]	[0.024]	[0.028]	[0.023]	

Far from entrance					
Kernel	−0.0108	0.0394	0.0515*	0.1048***	0.0533**
Standard error	[0.069]	[0.040]	[0.030]	[0.024]	
$n = 4$, caliper = 0.001	0.1782*	0.0820*	0.1219	0.0519	−0.0700
Standard error	[0.093]	[0.042]	[0.093]	[0.042]	
$n = 4$, caliper = 0.01	−0.0112	0.0818*	0.0128	0.0617	0.0489
Standard error	[0.080]	[0.045]	[0.066]	[0.039]	
$n = 1$, caliper = 0.01	0.1097	0.1358	0.0492	0.0718	0.0225
Standard error	[0.094]	[0.112]	[0.092]	[0.089]	

Notes: *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. Standard errors in brackets. Bias adjustment.
[a] Test based on a seemingly unrelated estimation (suest); see Weesie (1999) and updated versions.

Table 6. *National parks' effects on workers' wages in areas close to the entrance, by sex and pre-treatment population density*

	Males	*Females*	*Difference*
Kernel	0.0647***	0.1646***	−0.0999**
	[0.021]	[0.041]	[0.046]
$n = 4$, caliper = 0.001	0.0723***	0.1918***	−0.1195*
	[0.028]	[0.061]	[0.066]
$n = 4$, caliper = 0.01	0.0676**	0.2018***	−0.1342**
	[0.027]	[0.050]	[0.056]
$n = 1$, caliper = 0.01	0.0626*	0.1552**	−0.0926
	[0.038]	[0.072]	[0.080]
	Low population density	*High population density*	*Difference*
Kernel	0.1710***	−0.0388	−0.2099***
	[0.030]	[0.052]	[0.060]
$n = 4$, caliper = 0.001	0.2007***	0.3681	0.1674
	[0.046]	[0.359]	[0.343]
$n = 4$, caliper = 0.01	0.1964***	0.0592	−0.1372
	[0.042]	[0.254]	[0.251]
$n = 1$, caliper = 0.01	0.2215***	0.1843	−0.0372
	[0.050]	[0.829]	[0.790]

Notes: *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. Standard errors in brackets.

high population density in 1973. We find that a large and significant effect is present in areas with low pre-treatment population density. Because low population density is associated with low development, this might suggest that establishing a national park has significant positive effects in less developed areas. The estimated effects for high population density are not robust. Therefore, the estimated differences in the effects between high and low population density areas are also not robust.

5.1. *Robustness tests*

As an alternative model, we test whether the continuous variable distance to national parks has an effect on wages. We use a restricted sample based on kernel matching and include as controls the same variables as in the core results, i.e., year effects, worker innate characteristics, geographic variables and pre-treatment characteristics. Qualitatively, we find that the core results hold (table 7, panel A). The effect is negative and significant for those living close to the park, i.e., as distance from the park increases, wages decrease. The distance coefficient is also negative for those living close to the entrance. Finally, the distance coefficient is insignificant when testing the effects far from the entrance.

Also as a robustness test, we test whether the choices of the distance cutoffs affect the results. For the core regressions, we used a 5 km ring around the national parks and a distance to the entrance up to 20 km by roads to define the treated observations. As robustness tests, we ran the core model

Table 7. *Robustness checks on the effects of the distance from the national parks on wages, and placebo test*

	(1) *Close to park*	(2) *Close to entrance*	(3) *Far from entrance*
Panel A: Effects usingcontinuous distance[a]			
Effect	−0.0425***	−0.0572***	0.0985
	[0.010]	[0.014]	[0.109]
Panel B: Effects by distances to National Parks			
0–4 km (linear)	0.0717***	0.1276***	−0.0630
	[0.026]	[0.028]	[0.060]
0–6 km (linear)	0.0575***	0.0677***	−0.0460
	[0.019]	[0.021]	[0.034]
0–18 km by roads		0.0501**	−0.0019
		[0.021]	[0.057]
0–22 km by roads		0.0858***	−0.0619
		[0.019]	[0.056]
Panel C: Placebo effect			
$n = 4$, caliper = 0.001	0.0111	0.0144	0.0339
	[0.015]	[0.017]	[0.054]
$n = 1$, caliper = 0.01	−0.0073	0.0168	0.0160
	[0.028]	[0.035]	[0.102]
$n = 4$, caliper = 0.01	0.0099	0.0005	0.0602
	[0.016]	[0.017]	[0.049]

Notes: Controls used in the models in this table: year effects, worker's characteristics, geographic characteristics, characteristics of 1973.
[a]Regression used same observations as in the previous evaluations of each treatment with weights based on kernel matching.
$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Standard errors in brackets.

for 0–4 km and 0–6 km buffers, and 0–18 and 0–22 km distance by roads from the entrance. We find similar results (table 7, panel B). The estimated impacts close to the entrance are positive and significant and the estimated impacts far from entrance are insignificant.

Finally, we conduct a placebo test. Following Canavire-Bacarreza and Hanauer (2013), we identify a placebo group using the matched controls from each of the analyses. We then rerun the matching analysis using the placebo group as treated observations and the rest as untreated observations. With this analysis, we are trying to show that our results are driven by the implementation of parks and not by something other than the park. If the implementation of the parks is what drives the result, we should not see any statistically significant effect in the placebo tests. This would imply that covariates are creating a quality counterfactual (Canvire-Bacarrreza and Hanauer, 2012). This is exactly what we found using three different matching strategies (table 7, panel C). The estimated placebo effects are statistically insignificant for all three matching strategies tested.

6. Conclusions and discussion

We estimated the effect of national parks on local workers' wages by comparing those close to parks with workers living in similar areas far from parks. We found that there are positive effects of national parks on wages, but these effects vary according to location and characteristics of the individuals. There is strong evidence that workers close to a park's entrance benefit from the park's establishment. Protected areas can generate benefits, especially when accompanied by tourism development. These results are robust to the model specification when using the explanatory variable of interest as categorical or continuous, and to different distance buffers around the parks.

Additionally, we used individual-level data. Therefore, we were able to separate those who were born in the community from those who migrated from somewhere else. We showed that, when considering only local workers, the positive effects are significantly lower than when considering both local and immigrant workers. Previous analyses that use aggregated observations might have overstated the magnitude of the benefits of protected areas for local people.

Our data also allowed us to test heterogeneous effects using individual characteristics. We found that women receive a higher premium from being close to park entrances than do men. This result is consistent with switching from agriculturally intensive activities to service activities. Moreover, we found that parks have strong effects in areas with low initial population density. This suggests that parks might have large benefits in less developed areas, which are associated with lower population densities.

In conclusion, our results suggest that, even if there are socioeconomic benefits on average (higher wages), local people are affected differently depending on the access they have to tourism-related activities. We showed that one of the possible mechanisms through which protected areas benefit local people is indeed tourism. Therefore, policies that encourage tourism with adequate regulations so that conservation is not threatened might increase local people's benefits.

There are two important issues in relation to our results that are important to mention. First, if wages are higher around the parks, one could expect that prices of other goods will be higher. If this is the case, there might not be benefits for those living close to the entrances of the parks in real terms. However, there is evidence that the main determinants of prices for locals in areas around the park are related to transport costs and macroeconomic factors that affect all rural areas (Villalobos, 2009). Still, it is an important issue that should be considered when promoting tourism.

Second, there are other types of protected areas that have not been analyzed in this paper. Some of them are stricter in terms of visitors. This limits the potential benefits from tourism. By analyzing the treatment 'far away from entrance', we shed some light on this issue. However, it is not clear if those protected areas will have a different dynamic because they do not have an entrance at all. Future research could focus on learning about this type of protected area.

Along these lines, more detailed analysis is required to better understand the effects of national parks on local communities' welfare. For instance, the tourism effect might be looked at more closely by testing whether those close to the most visited national parks benefit more. Also, panel data analysis might help better estimate dynamic effects. It would also be important to test other dimensions of welfare such as education, infrastructure and migration.

References

Andam, K., P.J. Ferraro, K.R. Simis, A. Healy, and M. Holland (2010), 'Protected areas reduced poverty in Costa Rica and Thailand', *Proceedings of the National Academy of Sciences* **107**(22): 9996–10001.

Black, S.E. and F.E. Strahan (2001), 'The division of spoils: rent-sharing and discrimination in a regulated industry', *American Economic Review* **91**(4): 814–831.

Blanchflower, D.G., A.J. Oswald, and P. Sanfey (1996), 'Wages, profits and rent-sharing', *Quarterly Journal of Economics* **111**: 227–251.

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

Dehejia, R.H. and S. Wahba (2001), 'Propensity score-matching methods for nonexperimental causal studies', *Review of Economics and Statistics* **84**(1): 151–161.

Duffy-Deno, K.T. (1998), 'The effect of federal wilderness on county growth in the intermountain western United States', *Journal of Regional Science* **38**(1): 109–136.

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

Fortin, M.J. and C. Gagnon (1999), 'An assessment of social impacts of national parks on communities in Quebec, Canada', *Environmental Conservation* **26**(3): 200–211.

Funkhouser, E. (1996), 'The urban informal sector in Central America: household survey evidence', *World Development* **24**(11): 1737–1751.

Funkhouser, E. (1998), 'Changes in the returns to education in Costa Rica', *Journal of Development Economics* **57**(2): 289–317.

Heckman, J. (1974), 'Shadow prices, market wages, and labor supply', *Econometrica* **42**(4): 679–694.

Heckman, J. (1979), 'Sample selection bias as a specification error', *Econometrica* **47**(1): 153–161.

Hill, J.L., J. Brooks-Gunn, and J. Waldfogel (2003), 'Sustained effects of high participation in an early intervention for low-birth-weight premature infants', *Developmental Psychology* **39**(4): 730–744.

Ho, D., K. Imai, G. King, and E. Stuart (2007), 'Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference', *Political Analysis* **15**(3): 199–236.

IUCN (1994), *Guidelines for Protected Areas Management Categories*, Cambridge and Gland: International Union for the Conservation of Nature.

Joppa, L.N., S.R. Loarie, and S.L. Pimm (2009), 'On population growth near protected areas', *PLoS ONE* **4**(1): e4279.

Lewis, D.J., G.L. Hunt, and A.J. Plantinga (2002), 'Public conservation land and employment growth in the Northern Forest Region', *Land Economics* **78**(2): 245–259.

Lewis, D.J., G.L. Hunt, and A.J. Plantinga (2003), 'Does public lands policy affect local wage growth?', *Growth and Change* **34**(1): 64–86.

https://doi.org/10.1017/S1355770X14000461 Published online by Cambridge University Press

List, J., M. Margolis, and D. Osgood (2006), 'Is the Endangered Species Act endangering species?', NBER Working Paper No. 12777, National Bureau of Economic Research, Cambridge, MA.

Morgan, S.L. and C. Winship (2007), *Counterfactuals and Causal Inference: Methods and Principles for Social Research*, Cambridge: Cambridge University Press.

Mullan, K., A. Kontoleon, T.M. Swanson, and S. Zha (2009), 'Evaluation of the impact of the natural forest protection program on rural household livelihoods', *Environmental Management* **45**(3): 513–525.

Pfaff, A. and J.A. Robalino (2008), 'Human choices affect conservation impact: correct impact evaluation after the fact and looking forward', in C. Palmer and S. Engel (eds), *Avoided Deforestation: Prospects for Mitigating Climate Change*, Oxford: Routledge, pp. 192–207.

Pfaff, A. and J.A. Robalino (2012), 'Protecting forests, biodiversity, and the climate: predicting policy impact to improve policy choice', *Oxford Review of Economic Policy* **28**(1): 164–179.

Pfaff, A., J.A. Robalino, G.A. Sanchez-Azofeifa, K. Andam, and P. Ferraro (2009), 'Park location affects forest protection: land characteristics cause differences in park impacts across Costa Rica', *B.E. Journal of Economic Analysis and Policy* **9**(2): article 5.

Pfeffer, M.J., J.W. Schelhas, and L.A. Day (2001), 'Forest conservation, value conflict, and interest formation in a Honduran national park', *Rural Sociology* **66**: 382–402.

Robalino, J.A. (2007), 'Land conservation policies and income distribution: who bears the burden of our environmental efforts?', *Environment and Development Economics* **12**(4): 521–533.

Rosenbaum, P.R. and D.B. Rubin (1983), 'The central role of the propensity score in observational studies for causal effects', *Biometrika* **70**: 41–55.

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

SINAC (2006), *El Sistema de Áreas Silvestres Protegidas de Costa Rica: Informe Nacional para el II Congreso Mesoamericano de Áreas Protegidas*, San José: Sistema Nacional de Áreas de Conservación.

SINAC (2007), *Áreas silvestres protegidas de Costa Rica: Informe nacional para presentar al II congreso latinoamericano de parques nacionales y otras áreas protegidas*, San José: Sistema Nacional de Áreas de Conservación.

UN (2012), *The Millennium Development Goals Report 2012*, New York: United Nations.

Villalobos, L. (2009), 'The effects of National Parks on local communities' wages and employment in Costa Rica', MSc thesis, CATIE (Centro Agronómico Tropical de Investigación y Enseñanza), Turrialba, Costa Rica.

von Wacher, T. and J.F. Schmieder (2009), 'Does wage persistence matter for employment fluctuations? Evidence from displaced workers', *American Economic Journal: Applied Economics* **2**(3): 1–21.

Weesie, J. (1999), 'Sg121: Seemingly unrelated estimation and the cluster-adjusted sandwich estimator', *Stata Technical Bulletin* **52**: 34–47; reprinted in *Stata Technical Bulletin Reprints***9**: 231–248, College Station, TX: Stata Press.

Wittemyer, G., P. Elsen, W.T. Bean, A.C.O. Burton, and J.S. Brashares (2008), 'Accelerated human population growth at protected area edges', *Science* **321**(5885): 123–126.