



Universidad de
San Andrés

UNIVERSIDAD DE SAN ANDRÉS

DEPARTAMENTO DE ECONOMÍA / POSGRADO

DOCTORADO EN ECONOMÍA

**ECONOMICS OF RURAL INFRASTRUCTURE:
EVALUATING THE PROVISION OF TRANSPORT AND
ELECTRICITY IN PERU**

- 1. ELECTORAL CYCLES AND RENEGOTIATION OF
TRANSPORT INFRASTRUCTURE CONCESSION
CONTRACTS.**
- 2. IMPACT OF RURAL ELECTRIFICATION ON
EDUCATION: A CASE STUDY FROM PERU.**
- 3. ROADS AND AGRICULTURE: IMPACTS OF
CONNECTIVITY IN PERU**

AUTOR: JULIO CÉSAR AGUIRRE MONTOYA

LEGAJO: 4667450

MENTOR: MARTÍN ROSSI

BUENOS AIRES, MARZO 2016

Julio C. Aguirre

Electoral Cycles and Renegotiation of Transport Infrastructure Concession Contracts

Research Center, Universidad del Pacífico (Peru)
Email: aguirre_jc@up.edu.pe;
telephone: +(511) 219-0100, ext. 2146; fax: +(511) 2190135.

Abstract: Using a unique governmental database from Peru, I find that electoral cycles significantly affect the hazard of renegotiation of transport infrastructure concession contracts. The effect of electoral cycles remains robust when applying a varied set of specifications. In addition, electoral cycles have positive marginal effects on the hazard of renegotiation of concession contracts between 3% and 6%. These results provide a perspective on the importance of enforcing institutional and regulatory supervisory activities of infrastructure concession contracts in order to minimize possible political influences on the contractual stability of these contracts.

Keywords: Electoral cycles, Concession Contracts, Renegotiations, Survival Analysis.

JEL Code: D72, D86, L51

July 2013

1. Introduction

In the 1980s, several Latin American countries implemented drastic political and economic reforms. One of the key elements of these reforms was the privatization and concession of the majority of public utilities. This led to the involvement of the private sector in the construction and operation of public utilities under governmental regulation through the supervision of concession contracts.

There is an extensive literature on renegotiation of concession contracts that assesses the determinants of renegotiation focusing on contractual clauses such as mechanisms of regulation, award criteria, number of bidders in the auction of the contract, etc., but there are few studies of the relationship between electoral cycles and renegotiation. This relationship is the topic of this paper in which I apply quarterly-times series and hazard models to assess the relationship between electoral cycles and renegotiation.

Using a database of 27 transport infrastructure concession contracts from Peru (Table 1) awarded since 1994, I find that the timing of national election (i.e. the electoral cycle) significantly affects the incidence (and re-incidence) of renegotiation of contracts. My results are robust to the inclusion of quarter periods within and out-of the electoral year period, and also to controlling for some contractual features (including the designing of the contracts and the manner in which they were presented).

My intention is to provide a perspective on the importance of the enforcement of institutional and regulatory supervision of transport infrastructure concession contracts in order to try to minimize possible political influences affecting contractual stability.

In the Peruvian case, the institutional framework in the field of transport infrastructure includes four principal public institutions. A public agency, in charge of the designing of the concession contracts, PROINVERSION, which coordinates with the Ministries of Transportation and Communications (MTC), and Economy and Finance (MEF) technical features regarding the specific transport infrastructure needs for the country (roads, airports, ports and railroads) and the economical feasibility (Public Private Partnerships, for instance). Then the contract is awarded to a private operator through a tendering process, administered also by PROINVERSIÓN. After signing the concession contract between the private operator and the state (MTC), the

role of the transport infrastructure regulator (OSITRAN) begins supervising the effective compliance of the contract.

Although the previous institutional framework belongs to the increasing deployment of independent regulators with regulatory contracts in different countries (given that contracts supervised by independent regulators can be more sustainable and that independent regulators can enjoy greater credibility -Eberhard, 2007), a detailed analysis of addenda of concession contracts in Peru shows that renegotiations have been growing with each government (8, 24 and 37 modifications during Fujimori, Toledo and Garcia administrations, respectively)¹. This suggests to me that there exist possible political influences affecting contractual stability, and a proxy variable to be considered is the election process itself.

For instance, the autonomy of the regulators is not so clear because every five years the Board of Director along with the election of a new President of the Republic. In particular, in a first stage, an academic institution runs a selection process based on technical and academic capabilities of the candidates. However, in a second stage, a short list of the best three candidates is proposed to the newly elected President of the Republic, who has the final decision.² In essence, my results indicate that during electoral periods the rate of renegotiation increases (Figures 1 and 2), suggesting a weakness in the autonomy of the regulatory body.

While it is obvious that no contract is perfect, it is essential to be aware of its imperfections due the existence of transaction costs (Coase 1937; Williamson, 1995) that could determine the need to renegotiate it.³ However, although the renegotiation of contracts may result in better contracts that may have a positive effect on social welfare, it is of concern that frequent changes to contracts instead go against social interests, such as extra rents for the firm (after signing the concession contract, the operator has incentives to offer more favorable conditions than those offered by a competitive bidder

¹ According to information provide by Legal Area of OSITRAN.

² For the 2011 presidential elections, two forerunner candidates asked the outgoing government to postpone the election of the Board of Director of Regulators until after the presidential election, to be one of them who choose the authority (El Comercio, 2011). In 2012, after two consecutive contests, it was elected President of the Board of Transportation regulator. The first best two candidates in the two first rounds were put aside despite they fulfill technical and academic qualifications necessary to such functions.

³ In addition, it is not possible for contracting parties to define ex ante or predict all the possible contingencies that may occur after the signing of a contract. Moreover, even if this were possible, it would be very expensive to explicitly write all such contingencies into a contract and to expect that courts or officials would be able to check all actions under each contingency in order to enforce them (Tirole 1999).

when the contract is auctioned with a view to conducting renegotiations in the future), electoral advantage for the government (promising benefits to voters) or mutual benefits (capture and/or corruption features).

According to the Peru's transport infrastructure regulator, OSITRAN, the main reasons which have determined amendments to the infrastructure concession contracts are: (i) "works", which represents 47% of the cases of the renegotiations, meaning procedures for recognition of work in progress, details of some technical aspects, and additional investments necessary to modify the construction of the infrastructure; (ii) "funding" and "assets of the concession", each one with 10%. The former means some limitations accessing to financial support by banks, which will make easy the execution of the works or the operation of the concession; the latter refers to features about the concessions' assets in terms of ownership and management; (iii) "budget difficulties", which represents 8% of the renegotiations (and only related to Public-Private Partnership highway concessions), caused by necessary upgrade of the initial budget (which was elaborated based on preliminary studies before the concession) necessary for additional works and new traffic flows conditions roads; (iv) "payments to the operator", and "extension of duration of the contract" represent 6% each one. While the former refers to the recognition of additional activities defined at first place by contract, the latter deals with the term of the contract; (v) "land expropriation" and "completion of the contract" controversies have 5% each one; (vi) arbitration controversies with 4%; (vii) "tariffs settings" and "retribution to the state" with 3% each one; and, (viii) "others" categories such as modifications to the first technical design, salaries, insurance and guarantees.

Repeated renegotiation of contracts is not beneficial from an institutional perspective (Guasch, 2004). Of the current 26 transport infrastructure concession contracts supervised by Peru's regulator authority (Table 1), OSITRAN, 20 have been modified (76%) with 78 addendums (up to April 2013) and an average time from award to first renegotiation of 1.82 years.⁴ According to the World Bank, Peru has the highest rate of renegotiation in the region (*El Comercio*, 2011).

In the following discussion, Section 2 provides the literature and theoretical framework considered; Section 3 discusses the data, methodology and results; and Section 4 presents the conclusions.

⁴ Guasch (2004) found that renegotiation incidence and average time until renegotiation in the transport sector in Latin America and the Caribbean was 55% and 3.12 years, respectively.

2. Theoretical framework and related literature

There are not complete contracts because of the difficult to accounting for all possible contingencies (writing specific clauses dealing with any unknown even would be too costly) (Hart and Moore, 1988; Tirole, 1999), contracting parties are not completely rational (agents –governments and concessionaires- often make mistakes and need a learning period to approach an optimal solution, also new governments might have different goals than previous ones, which means that contracts must make tradeoffs among these dissimilar objectives), and commitment by agents helps determine the nature of the contract and influences the probability of renegotiation (Guasch, 2004) (government cannot conclude expropriation lands process or concessionaries can anticipate the possibility of renegotiation and at the auction stage they submit a lower bid to increase their probability of winning the concession).

Given that no contract is perfect, it is essential to be aware of its imperfect nature due the presence of transaction costs (Coase, 1937; Williamson, 1995) and asymmetric information (Laffont and Maskin, 1982; Myerson, 1982), which are relevant concepts for a developing country with difficulties of enforcement.⁵

In particular, my paper is based on the regulation model from Guasch et al. (2006), which is also an extension of the basic model from Laffont and Tirole (1993) and Laffont (2003) to account for imperfect enforcement and renegotiation. Guasch et al. (2006) model arrives to a probability of renegotiation⁶, where: $Pr(\text{renegotiation}) = (1 - v)(1 - \pi(x^E))$, considering the right side is the government's "tolerance for renegotiation", which depends on x^E , the investment in enforcement which can be understood as the existence of an enforcement institution (a regulatory agency,). As a consequence, the enforcement institution can be constrained by politics and state capture, affecting the social welfare. Then, empirically, it should be expected that the probability of renegotiation would be affected by the results of recent elections (Guasch et al. 2008).

Using that model, Guasch et al. (2003, 2005, 2006 and 2008) and Guasch (2003 and 2004) find that conditions conducive to renegotiations included combinations of contract characteristics, regulatory environments, and economic shocks. Renegotiations

⁵ The 2014 World Development Report (World Bank, 2013) and The Global Competitiveness Report 2014 (from the Global Economic Forum) site Peru in a relatively low position regarding institutional requirements.

⁶ The formal and extended development of this model can be seen in Guasch et al. (2006).

were most likely to take place: during recessions or after monetary devaluations, after elections, when concessions were awarded before a proper regulatory agency was put in place, when they were regulated by a price cap mechanism, and when the contract included some type of minimum income guarantee. Also, the existence of investment was shown to reduce the incidence of renegotiation while a poor institutional environment (captured by an index of bureaucratic quality) increased renegotiations. Specifically, those authors apply a probit analysis and use a dummy variable to indicate if the year was an election year, in order to ensure a new administration will honor the contract signed by a previous administration. They found that electoral cycles had a marginal effect of 3%-5% on the probability of renegotiation.⁷

I consider that the literature on the determinants of renegotiations of concession contracts from a theoretical and/or empirical perspective is broad, but the relationship between election cycles and modification of concession contracts is reduced. My paper tries to contribute to this body of literature. To facilitate this, I have studied the Peruvian transport infrastructure concession contracts using quarterly-time series analysis and, in addition, I assessed how electoral timing affects the frequency of renegotiation of contracts using a monthly-time survival analysis.

My work also contributes to the discussion of how the relation between electoral cycles and renegotiation can be motivated more on political interests rather than technical features. According to Transparency International (2006):

...political corruption is often associated to the electoral cycle and relates primarily to election funding (...). Political corruption often manifests itself as lobbying or as extortion, but also through *revolving door* moves of senior politicians directly into high business positions, often in companies whose activities were previously regulated or controlled by the same politician.

Corruption is linked to public procurement during the stage of contract design, specifically, when contract renegotiation is permitted or performed introducing substantial changes in the bidding process.

In addition, I try to infer some observational consequences based on the fact that capture, corruption or populism interests during electoral periods can affect the stability of concession contracts. Focused on the French car park sector, De Brux et al. (2011),

⁷ The author uses a database of over more than 1000 concession contracts in various countries in Latin America and the Caribbean in the water, transport, energy and telecom sectors over the period 1985 – 2000.

analyzes the link between renewals and contracts renegotiations in the car park in the French sector. They argue that this impact depends on the type and characteristics of the renegotiation and the discretionary power that the authority maintains according to the type of contract signed (Delegate Management or Public Procurement Contracts). Thus, if the contracts are finishing and the parties are satisfied, the probability of renewal will be greater than the case where the parties are in the opposite situation. The study found, among others, the following findings: (i) the speed of the first renegotiation, understood as a proxy of private contractual opportunism (as interpreted by Guasch (2004)), decreases the likelihood of renewal. (ii) electoral cycle (the change of the current political administration in the last year of the contract) reduces the likelihood of renewal, which is explained by the change of authority involves changing the party and the existence of a potential breach.

Due to limited information on money transfers or public budget, mainly from the State to the Company, I do not analyze whether the effect of renegotiations implies welfare losses, which goes against the spirit of the regulation of utilities. Some authors analyze the relationship between electoral cycles and public budget. For instance, Schady (2000) uses province-level data on monthly expenditures, socioeconomic indicators, and electoral outcomes to assess political influences on the timing and geographic distribution of FONCODES⁸ expenditures in Peru between 1991 and 1995. Schady (2000) concludes that FONCODES projects were directed at provinces in which the marginal political effect of expenditures was likely to be the greatest. In addition, Mejía et al. (2008), explore the factors that explained budgetary allocations for road infrastructure in Colombia during the first Uribe administration. They found some evidence suggesting the predominance of political criteria and that technical criteria did not play a role in decisions on investment in road infrastructure. This supports the hypothesis that spending on road infrastructure is a populist measure since those municipalities in which Uribe-sponsored Consejos Comunales met received an additional \$14.1 million in comparison to municipalities where such meetings did not take place.

⁸ FONCODES, the Fund for Cooperation in Social Development, was created by President Alberto Fujimori in 1991.

3. Data, methodology and results

Taking into consideration the review of the literature on this subject published thus far, this paper attempts to show that electoral cycles affects the rate of renegotiation of concession contracts by using quarterly-time series and hazard analysis.

3.1. Quarter-yearly time series analysis

I use 75 observations (quarters-yearly) considering data from 27 contracts signed between July-September 1994 and January – March 2013 administered by the Peruvian regulator, OSITRAN.⁹ In order to test the hypothesis that the electoral timing affects the incidence (and re-incidence) of renegotiation, I estimate the following equation:

$$\text{Incidence of renegotiation}_t = \alpha_t + \gamma \text{Electoral Cycle}_t + \beta \text{GDP}_t + \theta \text{Quarter}_t + \varepsilon_t \quad (1)$$

where *Incidence of renegotiation* is measured as the ratio number of modified concession contracts in the quarter-yearly t to stock of concession contracts (modified and unmodified) supervised by OSITRAN in the quarter-yearly t ; γ is the parameter of interest for the dummy *Electoral Cycle* taking the value 1 if the quarter-yearly t is an electoral period (between the July – September quarter to the April – June quarter);¹⁰ *GDP* is the current quarter-yearly gross domestic product (percentage change) and ε is the error term. I also include four quarter-yearly dummies *within* the electoral period and two quarter-yearly dummies each one *out-of* the electoral period (one dummy before the July- September quarter and the other one after the April – June quarter), *Quarter* (Diagram 1). In addition, I regress the same explanatory variables on the dependent variable *Re-incidence of renegotiation*, which is measured as the number of concession contracts modified more than once in the quarter-yearly t to the stock of concession contracts already modified supervised by OSITRAN in the quarter-yearly t . Summary statistics are reported in Table 2).

The series used in these estimations are stationary according to standard unit root tests.¹¹ To deal with potential heteroscedasticity and serial correlation, I compute Newey-West robust standard errors.¹² In column (1) of Tables 3 and 4, I report the

⁹ The data set is available by request.

¹⁰ In Peru, the elections occur always on April (every five years). However, the campaigns of political parties usually begin half way through the previous year of the elections.

¹¹ According to Dickey-Fuller test statistic the null hypothesis that the series has unit root is rejected. All results mentioned but not shown are available from the author upon request.

¹² I set the lag truncation value at 4, rejecting the null hypothesis that the series are heteroskedastic and serial correlated.

estimates of the equation (1).¹³ The coefficient of the electoral cycle is positive and statistically significant. The value of the coefficient implies an increase of the incidence and re- incidence of renegotiation of 11.2% and 11.8%, respectively, during an electoral period. Also, in column (2) of Tables 3 and 4 I show that results are robust including the GDP as a control of the macroeconomic conjuncture. This can be considered as an indication of the importance of political considerations (Guasch et al., 2008).

Additional examinations

In order to ensure the results indeed do have a causal interpretation I explore many alternative specifications. In columns (3) to (6) of Tables 3 and 4 I include the quarter-yearly dummies *within* the electoral period renegotiation. The coefficient of the electoral cycle remains positive and statistically significant, except on the fourth quarter-yearly dummy (April to June), suggesting that the electoral quarter-yearly period, when the election process (on April) occurs, increases the incidence and re-incidence of renegotiation on 29.7% and 30.0%, respectively.

In addition, I evaluate if the effect of the electoral cycle extends beyond the electoral period, a quarter-yearly previous the beginning and a quarter-yearly after the ending of the electoral period (Diagram 1). In columns (7) and (8) of Tables 3 and 4, the coefficient on both quarter-yearly dummies each one *out-of* the electoral period (one dummy before the July- September quarter and the other one after the April – June quarter) are not statistically significant.

Finally, in column (9) of Tables 3 and 4, the simultaneous inclusion of controls and dummies reinforce the positive effect of the electoral cycle on the incidence and impact of re-contract renegotiation, 34.3% and 35.2%, respectively.

The results in column (6) of Tables 3 and 4 indicate that the correlation between electoral cycle and the incidence and re-incidence of renegotiation of concession contracts is significant only during the electoral period, mainly during the last quarter-yearly of that period when the election is in process. This numerical result could advocate for the theoretical result of Guasch et al. (2003), finding that political cycles are likely to influence the incidence of renegotiations, given the possibility that government accepts renegotiation of concession contracts to the extent to which its interests are aligned with those of the firm.

¹³ The results remain robust when I do not consider a road concession contract awarded in 1994, 4 years before the creation of the regulator (OSITRAN).

3.2. Hazard analysis

For the hazard of renegotiation, I use a survival analysis to model the time elapsed between the granting of a concession contract and its renegotiation. A nonnegative random variable T is defined here as the duration (or spell) between the granting of a contract and renegotiation. Let us define the probability that a concession contract that has occupied a (unmodified) state for a time t leaves it in the short interval of length dt after t . The probability that such a contract leaves the state within an interval dt at or after t is $P(t \leq T < t + dt | T \geq t)$, where the conditioning event that $T \geq t$ is just the event that the state is still occupied at t , that has not left before then. Dividing this probability by dt , I get the average probability of leaving per unit time period over a short time interval after t , and by considering this average over shorter and shorter intervals, the hazard function is:

$$\theta(t) = \lim_{dt \rightarrow 0} \frac{P(t \leq T < t + dt | T \geq t)}{dt} = \frac{f(t)}{1 - F(t)} = \frac{f(t)}{S(t)} \quad (2)$$

which is the instantaneous rate of leaving per unit time period at t .¹⁴

The expression $\theta(t)dt$ can be interpreted as the probability of exit from the unmodified state in the short interval of length dt after t , conditional on the state being occupied at t . It is also perfectly logical to include factoring the probability of exit in the short interval of length dt after t *without* the condition $T \geq t$, but this is a quite different concept from the hazard function. The hazard function gives the probability that a contract will be modified at its twentieth month of operation whereas the unconditional concept gives the probability that a concession contract will be modified at the twentieth month of operation. In terms of relative frequencies, for example, $\theta(20)dt$ gives the proportion of twenty-months-operating contract which is modified within dt at its twentieth month of operation. The unconditional concept gives the proportion of contracts which are modified within dt at the twentieth month of operation.

Let the duration distribution function be $P(T < t) = F(t)$, $t \geq 0$, at the point t , and the associated probability density function be $f(t) = dF(t) / dt$. $f(t)$ is sometimes known

¹⁴ Equation (1) can be understood as: $\theta(t) = Risk = \frac{P(Failure)}{P(Survival)}$, where Failure is the unconditional

probability that an event will occur, Survival is the probability that “up until now” the event has not yet occurred, and Risk is the conditional failure rate – given the event has not yet occurred, what are the chances that it will occur? [<http://www.iser.essex.ac.uk/resources/survival-analysis-with-stata-module-ec968>, date accessed: June 12, 2011].

as the unconditional failure rate. Survival analysis makes use of the complement to the cumulative distribution function and is written as $S(t) = 1 - F(t) = P(T \geq t)$. That is, the survival function that gives the contract's probability of survival remains unmodified to t , and also can be expressed as¹⁵:

$$S(t) = \exp\left\{-\int_0^t \theta(s) ds\right\} \quad (3)$$

and, from (1) and (2):

$$f(t) = \theta(t) \exp\left\{-\int_0^t \theta(s) ds\right\} \quad (4)$$

For a parametric estimation, I introduce regressors (time-invariant and time-variant covariates) as relevant characteristics (covariates) from the concession contracts data, into the hazard function. Thus, at t this is defined as being conditional on the value of X :

$$\theta(t; x) = \lim_{dt \rightarrow 0} \frac{P(t \leq T < t + dt \mid T \geq t, x)}{dt} = \frac{f(t; x)}{S(t; x)} = \exp\left\{-\int_0^t \theta(s; x) ds\right\} \quad (5)$$

with time-invariant-covariates, and,

$$\theta(t; X(t)) = \lim_{dt \rightarrow 0} \frac{P(t \leq T < t + dt \mid T \geq t, X(t + dt))}{dt} = \exp\left\{-\int_0^t \theta(s; X(s)) ds\right\} \quad (6)$$

with time-variant covariates.

Even if an underlying duration is properly viewed as being continuous, measurements are necessarily discrete. When measurements are fairly precise, it is rational to treat the durations as continuous random variables. But when the measurements are coarse - for example, monthly or even weekly - it can be important to account for the discreteness of the estimation (Wooldridge 2010).

Thus, for the parametric estimation, I estimate a discrete time model based on Jenkins's (1995 and 2005) "easy estimation" methods, applying a standard binary dependent variable model, in which for each concession contract there are as many data

¹⁵ For the detailed mathematical derivation of the hazard and survival functions in nonparametric and parametric estimations see Lancaster (1990) and Jenkins (2005).

rows as there are time intervals at risk of the event occurring for each contract. If concession contract i 's survival time is censored, the binary dependent variable is equal to 0 for all i 's spell months; if contract i 's survival time is not censored, the binary dependent variable is equal to 0 for all but the last of i 's spell months (month 1, ..., T_{i-1}) and equal to 1 for the last month (month T_i). Thus, I have a panel data set in which each cross section observation is a vector of binary responses with covariates. Because of the sequential nature of the data, time-varying covariates are easily introduced (Wooldridge 2010).

The discrete time models are estimated by maximum likelihood¹⁶ and, for this paper, I use a discrete time proportional hazard model named the “complementary log-log”. The hazard model with time-invariant can be written as $\theta(t; x) = k(x)\lambda_0(t)$ where $k(\cdot) > 0$ is a positive function of x and $\lambda_0(t) > 0$ is called the baseline hazard. This is common to all the concession contracts. Individual hazard functions differ proportionately based on the function $k(x)$ of observed covariates (explanatory variables described above). $k(\cdot)$ can be parameterized as $k(x) = \exp(x\beta)$, where β is the vector of parameters of my focus. Thus, $\log \theta(t; x) = z(t) = x\beta + \log \lambda_0(t)$ and β_j measures the semi elasticity of the hazard with respect to x_j .¹⁷ I use the complementary log-log (“cloglog”) discrete time hazard function, $p(t)$, where

$$\log[-\log(1 - p(t))] = z(t) \Rightarrow p(t) = 1 - \exp[-\exp(z(t))] \quad (7)$$

The cloglog has the property that the resulting model is a discrete-time counterpart of an underlying continuous-time proportional hazard model (Prentice and Gloecker 1978).

The available dataset for the hazard exercise includes 26 contracts data - 3 airports, 14 roads, 3 railways and 6 ports - regulated and supervised by OSITRAN and granted between 2003 and 2011 (Table 1), involving 522 observations (right censoring)¹⁸.

Regarding the explanatory variables, in order to capture the relationship between electoral cycles and the hazard of renegotiation, I use a dummy variable which takes the

¹⁶ See Jenkins (1995) and Wooldridge (2010) for the construction of the likelihood function.

¹⁷ If x_j is the log of an underlying variable, say $x_j = \log(z_j)$, β_j is the elasticity of the hazard with respect to z_j .

¹⁸ This means that, at the time of observation, in those cases when renegotiations had not yet occurred, the total length of time between entry and exit from the state is unknown.

value of 1 if the month belongs to an electoral period and 0, otherwise. I expect find a positive relationship.

In addition, I consider as covariates other explanatory variables (some of them similar to those used by Guasch (2004) in his probit analysis), given that many renegotiations in short periods of time could be also explained by: (i) the designing stage of the contract: the mechanism of regulation of the transport infrastructure (a dummy variable 1 if a price cap scheme with RPI-X is considered and 0, otherwise) and the percentage of land area allocated to the operator (a dummy variable 1 if the 100% of land was allocated to the operator before subscription of the contract and 0, otherwise); (ii) the awarding stage of the concession process: the number of firms competing in the auction and the awarding criteria (taking 1 if the selection criteria implied an economic transfer from the firm to the government and 0, otherwise); and, (iii) institutional and economic conjectural factors: the experience and the squared experience of the regulator measured in years at the month of renegotiation and the monthly gross domestic production percentage change (to account for a possible impact of economic cycles). In addition, I introduce dummy transport sectors (railways, airport, port and roads).

I expect that the hazard of renegotiations increases when the mechanism of regulation does not consider productivity gains (price cap without RPI-X),¹⁹ and when lands were not fully granted. Summary statistics on these variables show that most frequently renegotiations have been mainly placed in the case of contracts with price-caps without RPI-X schemes and the percentage of allocated lands to the operator is below 100% (Table 5). In addition, I expect a negative impact of *bidders* because a competitive process (more than one bidder) gives the contract to an efficient firm²⁰, reducing the speed of a renegotiation), and a positive impact when the award *criteria* implies an economic transfer, because the firm could bid a higher offer (the winner curse) that cannot fulfill in the future. The descriptive data shows that most frequent renegotiations have occurred mainly when the solicitation is not competitive as characterized by fewer bidders pursuing the concession contract and the criteria of selection in the process implied a monetary transfer (lowest subsidy and/or a canon as the highest periodical payment from the concessionaire to the government).

¹⁹ Regarding this variable, Guasch (2004) considers price cap vs. rate of return (like a proxy of allocation of risks) rather than a price cap with or without RPI-X schemes. In that case, Guasch found that the probability of renegotiation increases when a price cap mechanism is used.

²⁰ Instead of number of bidders, Guasch (2004) considers the “type of award criteria”: competitive process vs. noncompetitive process (direct adjudication or bilateral negotiations). The author found that competitive process increases the probability of renegotiations.

Finally, I expect that the more regulators push to limit renegotiation the greater the impact on the monthly gross domestic production percentage change which accounts for a possible impact of the economic cycle on the hazard of renegotiation²¹.

In Table 6 I report estimates of the discrete time complementary log-log model (equation (7)) considering transport sector dummies. In column (1) the coefficient of the electoral cycles indicates that during an electoral period the hazard of renegotiation of transport concession contracts can increase 5.23%.

In the remaining columns in Table 6 I show that the effect of electoral cycle with alternative specifications increases the hazard of renegotiation of contracts in 3.3% to 5.3%. The magnitudes are similar to those found by Guasch (2004). Indeed, results remain unchanged when I either consider the mechanism of regulation or the percentage of land area allocated to the operator before subscription of the contract (columns (2), (3) and (4)), or the number of firms competing in the auction and the awarding criteria (columns (5), (6) and (7)), or the experience and squared experience of the regulator and monthly gross domestic production percentage change (columns (8), (9), (10) and (11)). In addition to the electoral cycle, the mechanism of regulation, the percentage of land allocated to the firm before subscription of the contract, the number of bidders, the award criteria and the regulator's experience have statistically significant effects on the hazard of renegotiation of transport concession contracts.

4. Conclusion and discussion

I apply a quarter-yearly time series and duration analysis on a database of 26 transport infrastructure concession contracts to identify whether electoral cycles have an effect on the renegotiation of these contracts.

I find the incidence and re-incidence of renegotiation of concession contracts increases during election periods. Results are robust to the inclusion of coefficients of quarter-yearly dummies *within* the electoral period renegotiation and quarter-yearly dummies *out-of* the electoral period (one dummy before the July- September quarter and the other one after the April – June quarter). I also run apply survival analysis that provides additional reassurance that the electoral cycle has an effect on the hazard of renegotiation of concession contracts. The results are statistically significant when I consider covariates such as the mechanism of regulation, the percentage of land

²¹ Guasch (2004) considers both the rate of growth of GDP and the evolution of the real exchange rate as a proxy for macroeconomic shocks. I have only considered the former.

allocated to the firm before subscription of the contract, the number of bidders, the award criteria and the regulator's experience.

Following to Montesinos and Saavedra (2012), the previous reveals the natural imperfection of concession contracts and provides important criteria for the future designing of contracts, which imply that intrinsic (mechanism of regulation, the percentage of land allocated to the firm before subscription of the contract, the number of bidders, the award criteria) and extrinsic (the regulator's experience, the economic environment, the election process) aspects of contracts should be analyzed carefully in order to not affect the stability of the contract.

Some observational consequences can be inferred from the results: renegotiation can be influenced by more electoral votes for a certain candidate or party (or populism) and/or by corruption or capture by interest groups. Then, there is an important political risk to mitigate that depends crucially on the extent to which government interests are aligned with those of the private operator. A more detailed analysis of this aspect would need to consider the nature of political changes. In particular, asymmetries might appear depending on whether the previous government cares more or less for the rents of the firm than its successor (Aubert and Laffont, 2004).

However, even I need more accurate data to prove these potential underlying reasons, I consider that the results suggest some weaknesses or vulnerability in the institutional strength of Peru's regulatory framework that leaves the door open to political influences and special interest group pressures, which may ultimately affect the welfare of society as a result of the temporary instability of negotiated contracts.

When election cycles exert influence on the renegotiation of contracts, it may ultimately affect the expected results of *ex-ante* competition through the granting of the contracts, implementing a strategy to generate rents *ex-post*. As a consequence, a policy measure that avoid any kind of political influence on technical decisions of the regulatory bodies can be that the elections of Presidents of Directors of these regulatory bodies has been performed during a phase that does not coincide with the period of presidential elections of the country. A good reference is provided by Gabillon and Martimort (2004) on how the design of credible monetary institutions ensures greater economic stability. Under the principal agent literature, the inflationary bias due to the time inconsistency of monetary policy can be avoided by delegating through contract the implementation of this policy to a central banker who is separated from the main government body. This physical separation is often viewed as an important ingredient of

the Central Bank's independence from the political sphere since it certainly insulates somewhat monetary policy from the day-to-day influence of political authorities and from fluctuations in the preferences of those political principals.

Finally, I cannot empirically prove the external validity of my results, but rather intuitively, by inferring what the great body of evidence suggests. As I mentioned before, Guasch (2004) and Guasch et al. (2003) are the closest references, regarding the empirical estimation of the main determinants of renegotiation of concession contracts in Latin America and The Caribbean. According to these papers, the renegotiation of concession contracts has been higher especially in the sectors of water and sanitation (74.4%), transportation (54.7%) and energy (9.7%). These conclusions are mainly based on the countries of Mexico, Brazil, Argentina, Colombia, Chile, Dominican Republic, Peru, Honduras, Guatemala, Venezuela, Trinidad and Tobago, Costa Rica, Panama and Uruguay, which together represent almost 30% of all concession contracts in region having been modified. Given that the election process is a determinant factor of the renegotiation of concession contract, it creates an interesting opportunity for future research in the controlling of the political parties.

Acknowledgements

I would like to thank Martin Rossi for his useful comments and suggestions and acknowledge the invaluable help of Francisco Taquiri and Gian Carlos Silva. Also, special acknowledgement goes to CESifo Group Munich's sponsorship for presentation of this paper at the CESifo Venice Summer Institute workshop on "The Economics of Infrastructure Provisioning: The (Changing) Role of the State" on July 2013.

References

Aubert, C. and Laffont, J-J (2004). Political renegotiation of regulatory contracts, *mimeo*.

Coase, R. (1937). The Nature of the Firm. *Economica*,
doi:10.1111/j.1468-0335.1937.tb00002.x

De Brux, J., Beuve, J. & Saussier, S. (2011). Renegotiations and Contract Renewals in Public Private Agreements - An Empirical Analysis.
<http://chaire-eppp.org/EPPP-annual-conference-2011/program>
Accessed 4 April 2011.

Eberhard, A. (2007). Infrastructure Regulation in Developing Countries. An Exploration of Hybrid and Transitional Models. Working Paper N° 4, Public-Private Infrastructure Advisory Facility.

El Comercio (Lima), May 3, 2011, B-1.

El Comercio (Lima), June 3, 2011, B-13.

Fierro, C. (2011). Concesiones y terrenos no saneados: cuando una obligación pierde su naturaleza. *Revista de Regulación en Infraestructura de Transporte* (6), 120-134.

http://www.ositran.gob.pe/repositorioaps/0/0/jer/presentacion_revista/Revista%206%20completa.pdf#page=122

Accessed 2 April 2011.

Gabillon and Martimort (2004). The benefits of central bank's political independence. *European Economic Review* 48, 353 – 378. Elsevier.

Guasch, J.L. (2003). Infrastructure Concessions in Latin America and the Caribbean: The Renegotiation Issue and Its Determinants. *Infrastructure and Financial Markets Review* 9(2), 1-6.

Guasch, J.L. (2004). *Granting and Renegotiating Infrastructure Concessions. Doing it Right*. WBI Development Studies, The World Bank, Washington, D.C.

<http://crgp.stanford.edu/events/presentations/gcr2/Guasch3.pdf>

Accessed 2 May 2011.

Guasch, J.L., Laffont, J-J, & Straub, S. (2003). Renegotiation of Concession Contracts in Latin America. Policy Research Working Paper (3011). Washington, D.C.: The World Bank.

doi:10.1596/1813-9450-3011

Guasch, J.L., Laffont, J-J., & Straub, S. (2006). Renegotiation of Concession Contracts: A Theoretical Approach. *Review of Industrial Organization*. 29(1), 55-73.

Guasch, J.L., Laffont, J-J., and S. Straub (2007). Concessions of Infrastructure in Latin America: Government-Led Renegotiation. *Journal of Applied Econometrics*. 22(7), 1267-1294.

Guasch, J.L., Laffont, J-J., & Straub, S. (2008). Renegotiation of Concession Contracts in Latin America: Evidence from the water and transport sectors. *International Journal of Industrial Organization* 26(2), 421-442.

Hart, O. and Moore, J. (1988). "Incomplete Contracts and Renegotiation", *Econometrica*, 56, 755-785.

Jenkins, S. (1995). Easy Estimation Methods for Discrete-Time Duration Models. *Oxford Bulletin of Economics and Statistics*. 57(1), 129-136.

doi:10.1111/j.1468-0084.1995.tb00031.x

Jenkins, S. (2005). Survival Analysis with Stata: Module EC968. University of Essex Summer School.

<http://www.iser.essex.ac.uk/survival-analysis>.

Laffont, J.J. y Maskin, E. (1982). The Theory of Incentives: An Overview, in *Advances in Economic Theory*, ed. by W. Hildenbrand. Cambridge: Cambridge University Press.

Laffont, J.J. and Tirole, J. (1993). *A Theory of Incentives in Procurement and Regulation*. Third Printing, Cambridge, United States: MIT Press.

Lancaster, T. (1990). *The Econometric Analysis of Transition Data*. Cambridge: Cambridge University Press, 6-8.

Mejía, L. B., Botero, F. & Rodríguez, J.C. (2008). Paving roads with votes? Budgetary Allocations for Road Infrastructure Projects in Colombia, 2002 – 2006. *Colombia Internacional* (68), 14-42.

Myerson, R. (1982): Optimal Coordination Mechanisms in Generalized Principal-Agent Problems. *Journal of Mathematical Economics*, 10, 67-81.

Montesinos J. and Saavedra, E. (2012). Algunos alcances en torno a la institucionalidad y renegociación de concesiones en la infraestructura de transporte de uso público en Perú. Universidad Alberto Hurtado, *mimeo*.

Prentice, R.L., & Gloeckler, L.A. (1978). Regression analysis of grouped survival data with application to breast cancer data. *Biometrics* (34), 57-67. Cited in Jenkins (1995).

Schady, N. R. (2000). The Political Economy of Expenditures by the Peruvian Social Fund (FONCODES), 1991-95. *The American Political Science Review* 94 (2), 289-304.

Tirole, J. (1999). Incomplete Contracts: Where Do We Stand? *Econometrica* (67), 741-781.

Transparency International (2006). Handbook for Curbing Corruption in Public Procurement.

http://archive.transparency.org/regional_pages/asia_pacific/newsroom/news_archive2/handbook_for_curbing_corruption_in_public_procurement_2006
Accessed 05 May 2011.

Williamson, O. (1995). *The Economic Institutions of Capitalism*. New York, NY: The Free Press.

Wooldridge, J.M. (2010). *Econometric Analysis of Cross Section and Panel Data*. 2nd ed. Boston, MA: The MIT Press.

World Bank. 2013. *World Development Report 2014: Risk and Opportunity—Managing Risk for Development*. Washington, DC: World Bank

WEF (2014). *Global Competitiveness Report 2014-2015. Insight Report*. World Economic Forum. Geneva.

Table 1. Concessions of Peruvian Transport Infrastructure Supervised by OSITRAN

Infrastructure	Year Signed
Airports	
1. Jorge Chávez International Airport	2001
2. First group of regional airports (Iquitos, Pucallpa, Trujillo, Tarapoto, Cajamarca, Tumbes, Anta, Chachapoyas, Talara, Piura, Chiclayo y Pisco)	2006
3. Second group of regional airports (Andahuaylas, Ayacucho, Juliaca, Arequipa, Puerto Maldonado y Tacna y Tacna)	2011
Roads	
4. CONCAR (Arequipa – Matarani)	1994
5. IIRSA-North (Multimodal Amazon North Axis: Paita – Yurimaguas)	2003
6. Road Network No. 5 (Ancón – Huacho – Pativilca)	2005
7. Road Network No. 6 (Puente Pucusana – Cerro Cerro Azul - Ica)	2009
8. Sun Highway (Trujillo – Sullana)	2007
9. IIRSA South T5 (Azángaro – Juliaca; Puerto de Matarani and Ilo)	2005
10. IIRSA South T4 (Azángaro – Inambari)	2005
11. IIRSA South T3 (Inambari – Iñapari)	2005
12. IIRSA South T2 (Urcos – Inambari)	2007
13. IIRSA South T1 (Marcona – Urcos)	2007
14. Buenos Aires – Canchaque	2009
15. Road Section (Mocupe – Cayaltí – Oyotún)	2009
16. Road Network 4 (Pativilca – Santa – Trujillo and Puerto Salaverry – Empalme R01N)	2009
17. Road Section: Chancay Roundabout/road/DV Pasamayo – Huaral – Acos	2010
18.- IIRSA Center, Section 2	2012
19.- Road Network DV Quilca, DV Arequipa (Reparticion) - DV Matarani - DV Moquegua - DV Ilo - Tacna - La Concordia*	
Railways	
20. Central Railway	1999
21. South and Southeast Railway	1999
22. Electric train – Line 1, Lima (Villa El Salvador – Av. Grau – San Juan de Lurigancho)	2011
Ports	
23. Matarani Port Terminal	1999
24. Muelle Sur New Container Port Terminal	2006
25. Paita Port Terminal	2009
26. Mineral Shipment Terminal	2011
27. Multipurpose North Terminal in the Port Terminal of Callao	2011
28. New Yurimaguas - Nueva Reforma Port Terminal	2011

Source: Online portal of OSITRAN: <http://www.ositran.gob.pe/0/home.aspx>.

* Granted but not signed to date.

Table 2. Summary statistics

	Mean	Standard deviation	Observations
Incidence of renegotiation ⁺	0.102	0.166	75
Re-incidence of renegotiation ⁺	0.107	0.168	75
Quarter-yearly gross domestic product (percentage change) ⁺	5.343	3.927	75

Notes: (+) Quarter-yearly data for transport infrastructure concession contracts of Peru, for the period Jul-Sep 1994 to Ene-Mar 2013.



Table 3. The Effects of Electoral Cycles on the incidence of renegotiation

	Dependant variable: incidence of renegotiation								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Electoral quarter	0.112** (0.0556)	0.108** (0.0524)	0.162** (0.0637)	0.112* (0.0667)	0.122* (0.0664)	0.0328 (0.0394)	0.105* (0.0529)	0.123** (0.0516)	0.343** (0.145)
GDP		0.00406 (0.00544)	0.00396 (0.00541)	0.0041 (0.00553)	0.0041 (0.00539)	0.00465 (0.00474)	0.00418 (0.00548)	0.00306 (0.00541)	0.00356 (0.00478)
D1 <i>within</i> electoral period			-0.219*** (0.0664)						-0.386** (0.148)
D2 <i>within</i> electoral period				-0.0187 (0.0975)					-0.236 (0.164)
D3 <i>within</i> electoral period					-0.0588 (0.0939)				-0.266 (0.162)
D4 <i>within</i> electoral period						0.297** (0.147)			
D. <i>ex ante out-of</i> electoral period							-0.045 (0.0276)		-0.029 (0.0265)
D. <i>ex post out-of</i> electoral period								0.202 (0.135)	0.2 (0.139)
Constant	0.0720*** (0.0158)	0.0513 (0.0339)	0.0519 (0.0338)	0.0512 (0.0344)	0.0511 (0.0338)	0.0484 (0.0312)	0.0532 (0.0343)	0.0417 (0.0299)	0.0409 (0.0273)
Observations	75	75	75	75	75	75	75	75	75
R-squared	0.09	0.099	0.188	0.1	0.106	0.262	0.102	0.174	0.368

Notes: Newey-West heteroskedasticity –and autocorrelation- consistent standard errors are in the parenthesis. All models are estimated by OLS. D1 to D4 and D. are dummy variables. *Significant at the 10% level; **Significant at the 5% level; ***Significant at the 1% level.

Table 4. The Effects of Electoral Cycles on the re-incidence of renegotiation

	Dependant variable: re-incidence of renegotiation								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Electoral quarter	0.118** (0.0561)	0.114** (0.0527)	0.170*** (0.0636)	0.117* (0.0662)	0.130* (0.0668)	0.0382 (0.0413)	0.111** (0.0532)	0.130** (0.0519)	0.352** (0.14)
GDP		0.0042 (0.00556)	0.00409 (0.00555)	0.00421 (0.00565)	0.00424 (0.0055)	0.00479 (0.00486)	0.0043 (0.0056)	0.00319 (0.00554)	0.00367 (0.0049)
D1 <i>within</i> electoral period			-0.226*** (0.0673)						-0.394*** (0.143)
D2 <i>within</i> electoral period				-0.0107 (0.103)					-0.232 (0.163)
D3 <i>within</i> electoral period					-0.0635 (0.0935)				-0.272* (0.157)
D4 <i>within</i> electoral period						0.300** (0.142)			
D. <i>ex ante out-of</i> electoral period							-0.0423 (0.0314)		-0.0261 (0.0309)
D. <i>ex post out-of</i> electoral period								0.204 (0.133)	0.202 (0.137)
Constant	0.0751*** (0.0161)	0.0538 (0.035)	0.0544 (0.035)	0.0538 (0.0354)	0.0536 (0.0348)	0.0508 (0.0323)	0.0556 (0.0354)	0.0441 (0.0312)	0.0433 (0.0288)
Observations	75	75	75	75	75	75	75	75	75
R-squared	0.098	0.107	0.199	0.107	0.114	0.269	0.11	0.181	0.376

Notes: Newey-West heteroskedasticity –and autocorrelation- consistent standard errors are in the parenthesis. All models are estimated by OLS. D1 to D4 and D. are dummy variables.*Significant at the 10% level; **Significant at the 5% level; ***Significant at the 1% level.

Summary statistics of explanatory variables

Variable		Modified	
		No	Yes
Mechanism of regulation	RPI-X		4 (23.53)
	No RPI-X	4 (100.00)	13 (76.47)
% of lands given to the operator	<100	4 (100.00)	11 (64.71)
	100		6 (35.29)
Bidders	1	1 (25.00)	11 (64.71)
	2		4 (23.53)
	3		2 (11.76)
	4	1 (25.00)	
	5	1 (25.00)	
	10	1 (25.00)	
Selection criteria	No transfer	4 (100.00)	3 (17.65)
	Transfer		14 (82.35)
Regulatory experience (years)	mean	12.25	7.53
Electoral	No		6 (35,29)
	Yes		11 (64,71)

Notes: The last column describes how many contracts were (and not) modified (only 21 concession contracts were modified between July 1999 and August 2012, from the 27 concession contracts). The sum of values in each block is equal to the number of observations. Percentages in relation to the total of a column are shown in parenthesis.

SanAndrés

Table 6. The Marginal Effects of Explanatory Variables on Hazard of Renegotiation

	Dependant variable: Hazard of renegotiation											
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
electoral	0.0523*** (0.0196)	0.0407** (0.0170)	0.0494** (0.0211)	0.00617** (0.00309)	0.0378** (0.0163)	0.0384** (0.0174)	0.0341** (0.0161)	0.0334* (0.0172)	0.0529*** (0.0201)	0.0338* (0.0176)	0.00393* (0.00212)	0.00119 (0.000901)
Mechanism of regulation land		-0.244 (0.180)		-1.000*** (6.24e-05)							-1.000*** (5.49e-05)	-1.000 (0.000)
			0.0525 (0.0509)	1.000*** (4.94e-05)							1.000*** (4.70e-05)	0.998*** (0.0230)
bidder					-0.00843*** (0.00308)		-0.00543 (0.00391)				-0.000760 (0.000702)	-0.000203 (0.000188)
criteria						0.0283** (0.0114)	0.0144 (0.0127)				0.00171 (0.00203)	-0.000154 (0.000377)
expreg								0.0103** (0.00523)		0.0105** (0.00472)		0.000597 (0.000517)
expreg2								-0.000696** (0.000335)		-0.000704** (0.000319)		-3.72e-05 (3.34e-05)
pibvar									-0.000544 (0.00184)	-6.70e-05 (0.000581)		
lnj	0.0143** (0.00622)	0.0153*** (0.00557)	0.0131** (0.00585)	0.00249*** (0.000875)	0.0108* (0.00562)	0.0135** (0.00651)	0.0113* (0.00598)	0.0120** (0.00605)	0.0143** (0.00618)	0.0120** (0.00608)	0.00194** (0.000874)	0.000552 (0.000571)
Dummies sector	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	522	522	522	522	522	522	522	522	522	522	522	522

Notes: Standard errors in parenthesis clustered at transport sector level (ports, airports, roads and railroads). The model estimated is the complementary log-log model. * Significant at the 10% level; ** Significant at the 5% level; *** Significant at the 1% level.

Figure 1: The Incidence of Renegotiation and Quarterly Electoral Cycles

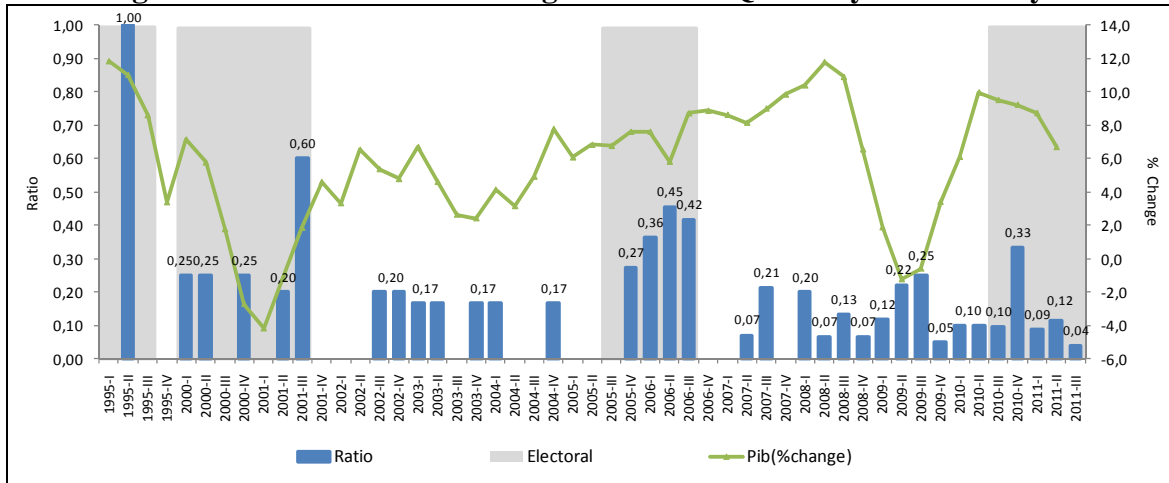


Figure 2: The Re-incidence of Renegotiation and Quarterly Electoral Cycles

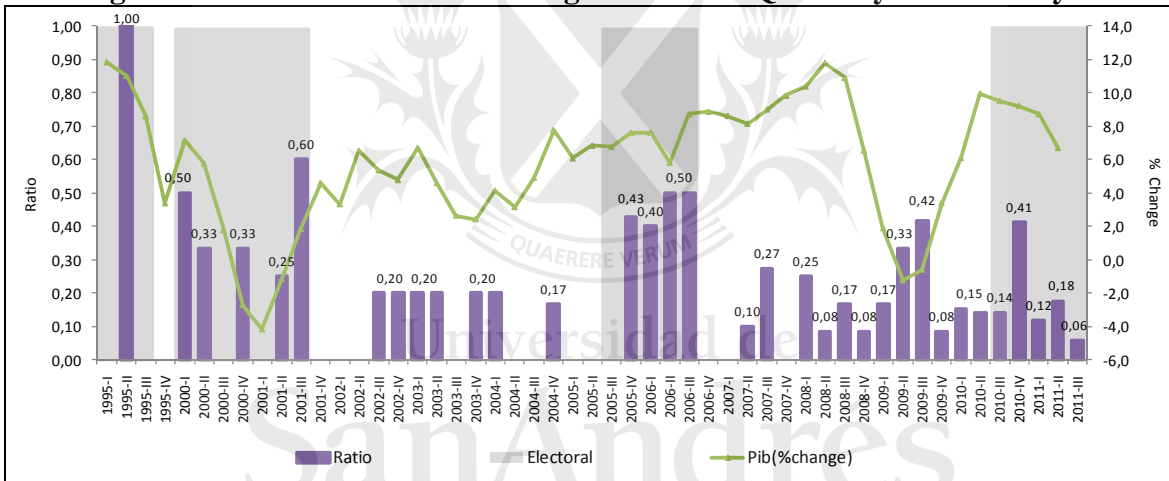
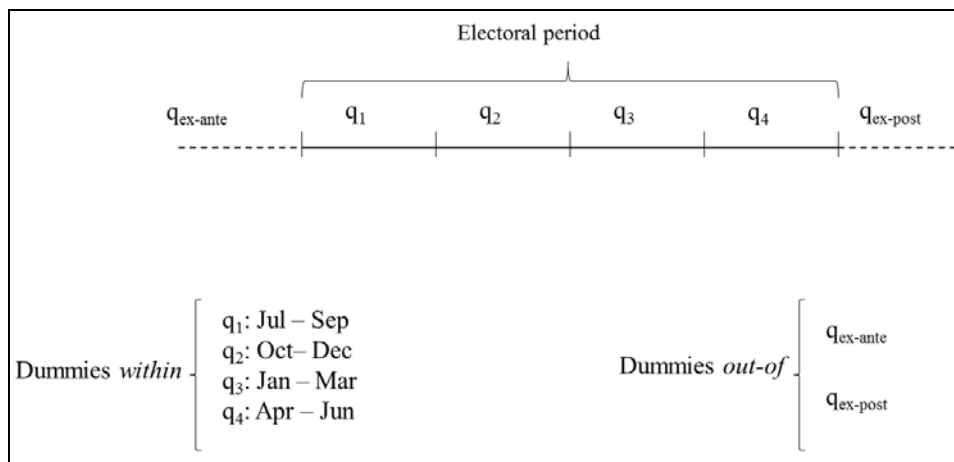


Diagram 1: Dummies within and out-of electoral period



**IMPACT OF RURAL ELECTRIFICATION ON EDUCATION:
A CASE STUDY FROM PERU**

Julio Aguirre*
Universidad del Pacífico

October 2014

In this paper, I study the impact of rural electrification on education. I find that connecting households to electricity increases studying time by children at home, and interpret this finding as indirect evidence of improvement in education. Using instrumental variables in order to overcome endogeneity problems, my results reveal that providing households with access to electricity leads to children studying an extra 35 minutes per day.

JEL Code: O12, C31, C81

Key words: Rural electrification; infrastructure; education.



Universidad de
San Andrés

*Research Center, Universidad del Pacífico (Peru) [aguirre_jc@up.edu.pe]. I thank Martin Rossi for useful comments and suggestions and acknowledge the invaluable help of Francisco Taquiri, Elmer Guerrero, Yohnny Campana and Lorenzo Oimas. Also, I am very grateful to the Ministry of Energy and Mines (MEM) and to the Energy Supervisor of Peru (OSINERGMIN) for the data and geo-spatial information.

I. Introduction

Impact evaluation of rural electrification programs has received considerable attention in previous literature, not only by academics but also by governments wishing to optimally allocate public resources. In both cases, the conclusion is the same: rural electrification generates substantial and favorable changes in welfare and it can be considered as a crucial prerequisite for economic growth (Khandker et al. 2012, 2013; ADB, 2010; Dinkelman 2008; IEG, 2008; ESMAP 2010, 2003; Cabraal et al. 2005; Martins, 2005; Barnes et al. 2003). The literature recognizes that the benefits of electricity can be divided into two categories: direct and indirect. The former includes improvements in lighting, television viewing, radio listening, and use of refrigeration; while the latter includes better educational outcomes, improved income-generation opportunities, lower fertility rates¹, and better health (by reducing indoor air pollution and the refrigeration of vaccines).

Rural electrification may affect education not only by improving the quality of schools resulting from their use of electricity-dependent equipment but also by increasing time allocation for studying at home (even though the availability of TV may decrease this time, it could also provide educational benefits) (IEG 2008). In this paper, I study the impact of rural electrification on education. I focus my analysis on the effects of connecting rural households to electricity on the time that children spend studying at home. Analyzing a unique survey on the use of energy in Peru conducted in 2013, I find that connecting households to electricity increases the time dedicated to studying by children at home and I interpret this finding as indirect evidence of improvement in education.

The evidence shows that children in electrified households have higher education levels than in those households without electricity. A survey applied in Peru in 2005 revealed that

¹ See Peters and Vance (2011).

children aged 6 to 18 in households with electricity and who attended school spent an average of 65 minutes per night reading and/or studying, whereas students in households without electricity spend 51 minutes on these activities. The survey found this difference to be statistically significant (ESMAP 2010). In a multicountry survey, IEG (2008), employed standardized demographic and health surveys in 9 countries – Bangladesh, Ghana, Indonesia, Morocco, Nepal, Nicaragua, Peru, Philippines and Senegal – and found that, on average, electricity increases the time that children spend studying at home by more than 70 minutes.

In Peru, through the Law of General Rural Electrification, a package of Rural Electrification Programs began in coordination with regional and local governments, providing villages (populated centers) and rural households with access to energy. The Ministry of Energy and Mines (MEM), through the Directorate General of Rural Electrification (DGER), established some prioritization criteria, most importantly lowest coefficients of provincial rural electrification, highest poverty rates, and amounts of subsidies required by connection and density of the population. Many Rural Electrification Programs were implemented between 1993 and 2013 and the last one was scheduled for the period 2008–2017.² To date, US\$ 278.6 million was invested, 5340 villages were connected to electricity, and the number of rural households with electricity increased from 7.7% in 1993 to 70% in 2013.³ At the end of 2012, DGER-MEM carried out a study to calculate the social benefits of rural electrification in order to: (i) determine whether public resources were being allocated efficiently, (ii) prioritize funding investments on

² It is important to say that Rural Electrification programs are deployed on villages in which household were already placed. In the Peruvian context, rural population use to set their households near to capital districts or where their crops are cultivated. Given that in the next paragraphs I will deal with the IV methodology, the distance between each population center and the nearest medium voltaje pylon is not (or it is difficult to be) a determinant of the decision to live in an specific place.

³ These projects include the installation of transmission lines of 60, 138 and 220 kV, 2872 km long, small-scale hydro and thermal generation units with 150 MW, 1523 solar panels. The implementation of the Rural Electrification Programs also included the extension of national grid networks and/or isolated electrical systems from which Rural Electric Systems are developed (MEM, Special Report from Diario El Comercio [Lima], December 15, 2013, and the National Plan for Rural Electrification 2013-2022).

public projects (to be implemented beginning in 2014), (iii) facilitate the social evaluation of investments on public projects in Peru, and (iv) estimate the direct and indirect benefits of rural electrification. The study lasted from November 2012 to March 2013 (Urrunaga et al. 2013). As part of that study, and in accordance with the requirements of DGER, the Rural Household Energy Use Survey (SRHEU) was conducted in February 2013.

The importance of the accountability in rural electrification programs has increased substantially through more frequent impact evaluation studies. As Ravallion (2008a and 2008b) documented, the methodological sophistication of some of these evaluations has increased substantially; however, he criticizes the dearth of rigorous evaluation research in development policies. In the field of rural electrification, extensive studies have been conducted to assess the impact of electrification by comparing connected and non-connected households within the same region (ESMAP 2003a; Madon and Oey-Gardiner 2002; Massé and Samaranayake 2002; World Bank 2006). Even though these studies have found that electrification provides significant benefits, most have simply shown that there is a correlation between rural electrification and development; however, this correlation could be due to a third variable, such as household income, and thus these studies do not necessarily demonstrate a causal relationship.⁴ For the purposes of this paper, the problem of endogeneity that exists in the implementation of rural electrification programs generates flaws in the determination of the direction of causality. In order to control for this problem, I use Instrumental Variables (IV) to determine the net effect of rural electrification, applied to cross-sectional data. A similar technique is used by Khandker et al. (2009, 2012) for education outcome variables.

⁴ In the case of other infrastructure (transport, for instance), the literature is also unsatisfactory, failing to address endogeneity issues convincingly (Straub 2013).

Khandker et al. (2009, 2012) estimate the benefits of rural electrification in Bangladesh and India, respectively, on various households and individual welfare outcomes (expenditure, income, energy consumption, employment, years of schooling, and time studying) using the 2005 Bangladesh Survey of Rural Households and the 2005 India Human Development Survey of Rural Households. In both cases, Instrumental Variables regression is used.⁵ In the former study, they use a household's location within or beyond 100 feet of an electrical line as an instrument since this influences household's adoption of grid electricity,⁶ but do not directly influence their outcomes. In the latter study, the proportion of households in a community who have electricity is expected to serve as an instrument because peer pressure or the demonstration effect is likely to affect a household's electrification decision since households tend to follow their neighbors in the village.⁷ They find that access to electrification increases weekly boys' and girls' time studying by more than 6 and 8 minutes/day, respectively, in Bangladesh, and by more than an hour in India (the increase is slightly higher for girls than boys). In this study, I use IV regression. The instrument is the topographic distance between each population center and the nearest medium voltage line, given that this variable is correlated with a household's connection status since the shorter the distance, the greater the likelihood of connection, but is not correlated with children's time studying at home. My results show a positive impact of household electrification on the time children spend studying.

⁵ In addition to Propensity Score Matching (PSM) for Bangladesh and Fixed Effects methods for India.

⁶ Considering that for those households living within 100 feet of the electricity line or lines that run through a village, connection cost for obtaining electric service is highly subsidized and therefore low (and, by contrast, the connection cost charged by the electric cooperatives for households beyond 100 feet of the line is much higher since they have to bear the full cost of connection),

⁷ If neighbours obtain electricity, then a household without electricity can signal lower socioeconomic standing, since households would be expected to avoid obtaining electricity if they could not afford it. It is expected that the higher the percentage of connected households in a village, the greater the likelihood that a household living in that village will connect to electricity, provided it can afford the connection fee and other associated costs. In addition, the proportion of village households with electricity should not directly impact a household's outcome.

This paper continues as follows: Section II describes the data and presents the econometric method; Section III reports the results; and Section IV presents the conclusions.

II. Data and methodology

In order to identify the effects of households' connection to electricity on education, I use the Survey of Rural Household Energy Use 2013 (SRHEU 2013) conducted by the Ministry of Energy and Mines of Peru (MEM). The survey included 987 electrified (654) and non-electrified (333) households in rural areas in 96 rural population centers⁸ in Peru. The sample was probabilistic, stratified at three stages: the levels of provinces, districts, and rural population centers.

The information collected includes comprehensive data on the socio-demographic characteristics of connected and non-connected households. This data includes household composition (size and each member's age, sex, and relationship to the head of household), demographics (education levels), economic indicators (assets, income, expenditures), and the way households use energy. Also, for each household member, individual measures of time use (e.g. hours of study at home) were collected. Table 1, columns (1) and (2), provides information on main household-level characteristics.

**** INSERT TABLE 1 ****

The effects of the provision of electricity to a region can be assessed using the conceptual framework of the theory of change (Bensch et al. 2011) in which the development project is typically represented in a results chain that links the intervention's input and activities (new electrification interventions in a region and households) to its outputs and impacts (translated into poverty reduction via different channels). Table 1, columns (3) to (6), assesses to what

⁸ The definition used by MEM for the purpose of the survey is that rural population centers are those with less than 100 dwellings grouped contiguously.

extent the comparability of household characteristics described above translates into heterogeneity between connected and non-connected households. The *p-values* presented in this table show that the tests for difference-in-means between the connected and non-connected households are significant for most of the characteristics.

I focus only on impacts of electricity connection on the time that (primary school) children use for studying at home. I employ this outcome as an intermediate measure to approximate the transmission channel to ultimate educational impacts. Descriptive statistics for this indicator are also provided at the bottom of Table 1. They show that a difference exists between connected and non-connected households at the national level.

Following Bensch et al. (2011), from an impact evaluation perspective, the survey that I use serves for an ex-ante impact assessment by comparing households already electrified to those not yet electrified using cross-sectional methods.

As mentioned previously, it has to be kept in mind that these kinds of intervention programs are difficult to evaluate, given the endogeneity problem that can arise. There is a potential self-selection process due to which comparing outcomes from connected and non-connected households may suffer from substantial biases (Ravallion 2008) because in the case of electrification interventions, the decision to connect is a choice of individual households, which may be taken for unobservable reasons and, at the same time, affect the outcome measured (Peters 2009). For instance, when using such a cross-sectional comparison, the impacts on children's studying time at home are difficult to evaluate. The reason for this is that households with parents with higher levels of education are more likely to raise the funds necessary to connect to the grid (because they have more income and know the importance of children studying at home). This simultaneity time-studying and connection status implies that it is not

possible to know if a household has parents with higher levels of education because it is connected, or if it is connected because they have higher levels of education.

Households with electricity have higher levels of benefits compared to those without a connection to electricity. The next question to ask is whether this means that having electricity in a household conclusively contributes to a better education for children in these households. Keeping in mind the endogeneity problem and that grid electricity service is extended first to more developed and densely populated regions (for revenue maximization), and only later reaches more remote and poorer areas (Khandker et al. 2009), I need to find a suitable instrument, that is, a variable that is correlated with electricity connection status but uncorrelated with the household's outcome variable. Other authors use the Propensity Score Matching (PSM) technique together with the IV method (Bensch et al. 2011; Khandker et al. 2009, 2001) on cross-sectional data. Certainly, both methods have their own advantages and disadvantages. An IV method controls for both observed and unobserved characteristics, while PSM cannot control for unobserved characteristics. But unlike IV or any regression technique, PSM does not assume a functional form, which is certainly an advantage (Khandker et al. 2009). Despite these reasons for using IV regression, I also test whether the OLS or the IV approach is the more appropriate estimation technique for the data at hand, using a Hausman test.

Here I use topographic distance between each population center and the nearest medium voltage line as the instrument for being connected to the electricity network. This measure was generated using Arcgis 10.1 software with the coordinates of the location of transmission lines provided by the Peruvian agency that supervises investment in energy and mines, OSINERGMIN. The Appendix shows the distance measures calculated. I consider this variable

to be correlated with household connection status since the smaller the distance, the greater the likelihood of connection, but not correlated with time studying at home by children.

III. Results

In this study, I am interested in estimating the causal effect of rural household electricity connection status on the time spent studying at home by children. Formally, I want to estimate the following equation:

(1)

$$\text{Time spent studying by children}_{ir} = \beta + \alpha \cdot \text{Household connection status}_{ir} + \theta \cdot X_{ir} + \delta_r + \varepsilon_{ir}$$

where *Time spent studying by children*_{ir} is the time (in hours) that children dedicate to studying at home *i* in region *r*; *X* is a vector of control variables; the δ_r is a region effect; α is the average treatment effect; and ε_{ipr} is an error term.

To address the endogeneity of household connection status on the time spent studying, I estimate equation (1) by Two Stage Least Squares (2SLS), where the endogenous dummy variable “household connection status” is instrumented by the exogenous variable “topographic distance”. Figure 1 plots the conditional probability of a household’s connection to electricity given the distance between the village and the nearest transmission line center. The most important feature of this figure is the negative relationship between the household’s probability of connection to electricity as the topographic (or linear) distance increases. First-stage estimates are reported in Table 2. The point estimates of the coefficient on topographic distance from the sample indicates that the probability of a household’s connection to electricity is around 18 percentage points higher, respectively, for those households located in villages nearest to transmission lines compared to those located far from transmission lines. First-stage effect is very precisely estimated and significantly different from zero.

** INSERT FIGURE 1 **

** INSERT TABLE 2 **

The IV estimator does not recover average treatment effects, unless I am willing to assume a constant treatment effect. Under sensible assumptions, however, it recovers an alternative parameter denoted Local Average Treatment Effect (LATE) (Angrist et al. 1996), which is the average effect of treatment on those individuals whose treatment status is induced to change by the instrument (i.e., by the dummy variable topographic distance). The households are compliers because they obtain a connection to electricity because they live in villages that are located near transmission lines, and would not have been connected otherwise. Thus, the results reported below do not need to be generalized to the population of households that under no circumstances would have been connected to electricity.

My estimates of the impact of households' connection to electricity are reported in Table 3, with and without head's household, household characteristics, and dummy regions. As a benchmark, I also report reduced-form estimates in columns (3) and (4). The preferred 2SLS estimates in column (6) indicate that connection to electricity significantly increases the time spent studying by children (58.40%).⁹ Thus, my instrumental variables results suggest that acquiring a connection to electricity allows children to study 35 more minutes. Children in connected households tend to study an additional 35 minutes per-day compare to children in non-connected households. The result of the Hausman test rejects the null hypothesis in these outcomes at 5 percent levels, confirming that IV estimation is the better estimator (consistent) for this sample. I report the OLS results for comparison.

** INSERT TABLE 3 **

⁹ Percent change is calculated as $100 * \text{Estimate} / \text{mean time studying by children in non-connected households}$: 58.40% ($=100 * 2.289 / 3.9198$).

IV. Conclusions

The purpose of this article is to contribute to the economic literature related to impact assessment of rural electrification programs. Using an IV approach to overcome endogeneity concerns, I find a positive association between rural electrification and the number of hours of study by school-age children, suggesting greater opportunities for improvement in school performance.

According to the findings of this study, electricity leads to a significant increase in the time children spend studying in rural households by 34 minutes. It is important to recognize that my results assume that all students graduate from primary school.

For the sake of motivating discussion, the benefits of rural electrification can be approximated in monetary terms: whether it is assumed that one hour of study by children between 3 and 12 years reduces the possibility of repeating the academic year by 1.6 percentage points (Beltrán 2013) and that the yearly-cost per student of rural public school for the government is US\$ 2070.7¹⁰⁻¹¹, let me suppose that one more hour of study by a child could avoid losing US\$ 33.13 (=2070.7 x 0.016) per year because of children have not repeating a year. Then, given that the benefits of a connection to electricity by households at the national level is 0.59 more hours (or 35 minutes) of study at home, in a year the benefits of rural households connecting to electricity amount to US\$ 19.5 (=0.59 x 33.13) per child. Of course, the other benefits of rural electrification, such as, illumination, radio and TV, and refrigeration¹² should be added to this amount

¹⁰ According to the Ministry of Education statistics: <http://escale.minedu.gob.pe/>

¹¹ Exchange rate of S/. 2.8 per US\$ 1.0.

¹² See Urrunaga et al. (2013), in which the costs necessary for the provision of rural electrification were calculated considering benefits of illumination and radio&TV (using consumer excedent and avoided costs methodologies) and of education (using matching techniques),

On the other hand, my cross-sectional analysis has a potential shortcoming: the survey I use was conducted during school holidays (December to March), biasing the answers about hours of study at home.

Finally, even though there are many challenges to be overcome, an expansion of access to electricity in a way that is equitable, both institutionally and financially, can have significant development benefits in rural Peru.

References

Angrist, J., G. Imbens, and D. Rubin (1996). Identification of Causal Effects Using Instrumental Variables. *Journal of the American Statistical Association* 91(434): 444-55.

ADB (2010). Asian Development Bank's Assistance for Rural Electrification in Bhutan: Does Electrification Improve the Quality of Rural Life? Impact Evaluation Study. Manila, Asian Development Bank.

Barnes, D. F., H. Peskin, and K. Fitzgerald (2003). The Benefits of Rural Electrification in India: Implications for Education, Household Lighting and Irrigation. Unpublished manuscript, South Asia Energy Infrastructure, World Bank.

Beltran, A. (2013). El tiempo de la familia es un recurso escaso: ¿cómo afecta su distribución el desempeño escolar? *Apuntes Revista de Ciencias Sociales* 72: 117-56.

Bensch, G., J. Kluve, and J. Peters (2011). Impacts of Rural Electrification in Rwanda. Discussion Paper No. 6195, Institute for the Study of Labor (IZA), Bonn, December.

Cabraal, R. A., D. F. Barnes, and S. G. Agarwal (2005). Productive Uses of Energy for Rural Development. *Annual Review of Environment and Resources* 30: 117-44.

ESMAP (2010). Peru: National Survey of Rural Household Energy Use. Energy Sector Management Assistant Program, Special Report 007/10, World Bank.

ESMAP (2003). Rural Electrification and Development in the Philippines: Measuring the Social and Economic Benefits. Energy Sector Management Assistant Program, Special Report 255/03, World Bank.

ESMAP (2000). Reducing the Cost of Grid Extension for Rural Electrification. Energy Sector Management Assistant Program, Special Report 227/00, World Bank.

IEG (2008). The Welfare Impact of Rural Electrification: A Reassessment of the Costs and Benefits. Washington, D.C., Independent Evaluation Group (IEG).

Khandker, S., D. Barnes, and H. Samad (2013). Welfare Impacts of Rural Electrification: A Panel Data Analysis from Vietnam. *Economic Development and Cultural Change* 61(3): 659-92.

Khandker, S., H. Samad, R. Ali, and D. Barnes (2012). Who Benefits Most from Rural Electrification? Evidence in India. Policy Research Working Paper 6095, World Bank.

Khandker, S., D. Barnes, and H. Samad (2009). Welfare Impacts of Rural Electrification. A Case Study from Bangladesh. Policy Research Working Paper 4859, World Bank.

Madon, G. and Oey-Gardiner, M. (2002). EnPoGen Study in Indonesia. *ENERGIA News* 5(3): 11-3.

Martins, J. (2005). The Impact Use of Energy Sources on the Quality of Life of Poor Communities. *Social Indicators Research* 72(3): 373-402.

Massé, R. and Samaranayake, R. M. (2002). EnPoGen Study in Sri Lanka. *ENERGIA News* 5(3): 14-6.

Peters, J. and Vance, C. (2011). Rural Electrification and Fertility – Evidence from Côte d'Ivoire. *Journal of Development Studies* 47 (5): 753-766.

Peters, J. (2009). Evaluating Rural Electrification Projects: Methodological Approaches. *Well-Being and Social Policy* 5(2): 25-40.

Ravallion, M. (2008). Evaluating Anti-Poverty Programs. In R.E. Evenson and T.P. Schultz, editors, *Handbook of Development Economics*, v. 4. Amsterdam, North-Holland.

Straub, S. (2013). Policy Lessons from the Recent Literature on Transport Infrastructure Development. Toulouse School of Economics, forthcoming.

Urrunaga, R., J.L. Bonifaz, J. Aguirre, G. Aragón, and O. Jara (2013). Beneficios sociales de la electrificación rural. Metodologías y estimaciones. Documento de Trabajo 91, Centro de Investigación, Universidad del Pacífico, Lima.

Table 1. Summary statistics

	Mean (1)	St. Dev. (2)	Non-C (3)	C (4)	Difference (5)	<i>p-value</i> (6)
<u>Household characteristics</u>						
Household has electricity (1=Yes, 0=No)	0.337	0.473				
Homeownership (1=Yes, 0=No)	0.853	0.354	0.8593 (0.3479)	0.8408 (0.3664)	0.0185	0.0000
Household has connection to water network (1=Yes, 0=No)	0.400	0.490	0.3149 (0.4649)	0.5676 (0.4962)	-0.0320	0.0000
Household has connection to sanitation network (1=Yes, 0=No)	0.091	0.288	0.0382 (0.1919)	0.1952 (0.3969)	-0.1569	0.0000
Household has concrete, wood or corrugated roof (1=Yes, 0=No)	0.633	0.482	0.5719 (0.4952)	0.7538 (0.4315)	-0.1819	0.0000
Household has concrete wall (1=Yes, 0=No)	0.082	0.275	0.0459 (0.2094)	0.1532 (0.3607)	-0.1073	0.0000
Household has concrete or hardwood floor (1=Yes, 0=No)	0.323	0.468	0.2982 (0.4578)	0.3724 (0.4842)	-0.0742	0.0184
Observations		987	654	333		
<u>Head of household's social characteristics</u>						
Time living in the populated center	22.867	17.791	22.7584 (18.2324)	23.0778 (16.9200)	-0.3194	0.7907
Education of head of household (years)	7.176	3.921	6.9931 (3.8297)	7.5360 (4.0759)	-0.5429	0.0396
Age of head of household (years)	45.409	15.013	44.2324 (14.9276)	47.7207 (14.9327)	-3.4883	0.0005
Sex of head of household (Male=1, Female=0)	0.881	0.323	0.8761 (0.3297)	0.8919 (0.3109)	-0.0157	0.4699
Household size	3.806	1.763	3.7431 (1.8183)	3.9309 (1.6441)	-0.1878	0.1136
Observations		987	654	333		
<u>Intermediate outcome</u>						
Children's time studying at home (in hours)	4.091	2.414998	3.9198 (2.3359)	4.3731 (2.4894)	-0.4533	0.0340
Observations.		542	337	205		

Note: For columns (4) and (5), standard deviations are in parentheses, Non-C: non-connected households; C: connected households.

Figure 1. Connection to electricity status of population centers and topographic distance to transmission lines

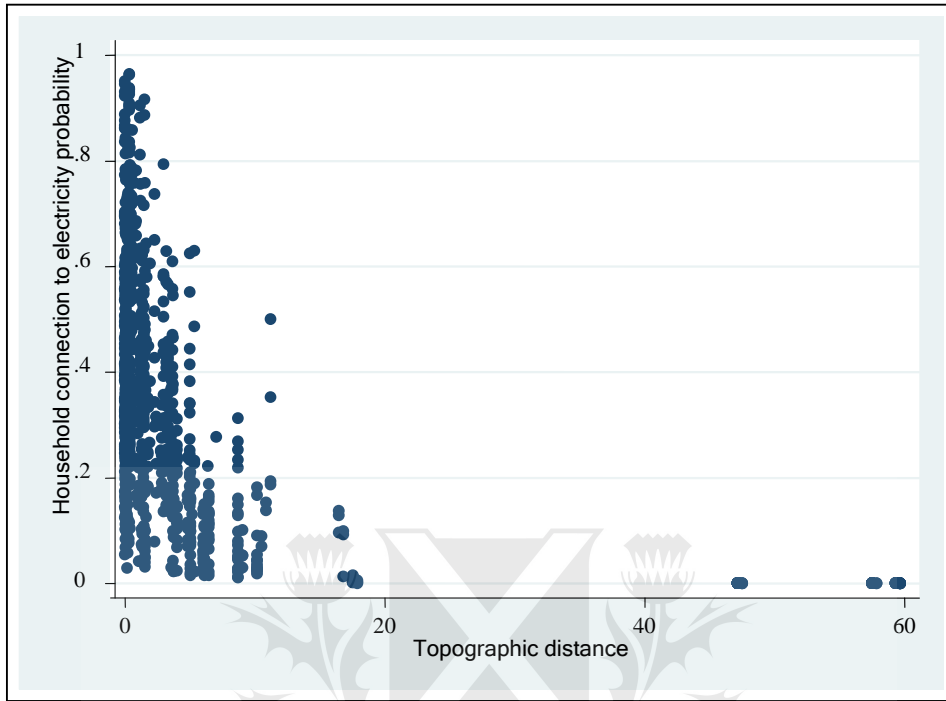


Table 2. First stage at a national level

	Dependant variable: household's connection to electricity	
	(1)	(2)
Topographic distance	-0.1830*** (0.040)	-0.1823** (0.0416)
Constant	-1.3257** (0.4387)	-1.7403* (0.4875)
Head's household and household characteristics	Yes	Yes
Dummy regions	No	Yes
Method	Probit	Probit
Observations	537	537
Wald test	$\chi^2(11)=83.99, \text{Prob}> \chi^2=0.000$	$\chi^2(13)=102.38, \text{Prob}> \chi^2=0.000$

Notes: Standard errors are in the parenthesis. *Significant at the 10% level; **Significant at the 5% level; ***Significant at the 1% level.



Universidad de
San Andrés

Table 3. Estimates of the impact of connection to electricity on children's time studying at home on the national level

	Dependant variable: children studying at home (hours)					
	(1)	(2)	(3)	(4)	(5)	(6)
Household's connection to electricity	0.399 (0.472)	0.739 (0.384)			1.561* (0.682)	2.289*** (0.564)
Topographic distance			-0.0246* (0.0099)	-0.0418*** (0.0101)		
Head's household and household characteristics	Yes	Yes	No	No	Yes	Yes
Dummy regions	No	Yes	No	Yes	No	Yes
Observations	537	537	537	537	537	537
Method	OLS	OLS	OLS	OLS	2SLS	2SLS

Notes: Standard errors are in the parenthesis. *Significant at the 10% level; **Significant at the 5% level; ***Significant at the 1% level. Models in columns (5) and (6) use topographic distance as an instrument variable. The Hausman test for endogeneity was applied. In all cases the null hypothesis was rejected at 5% levels, confirming that 2SLS estimations is a better method [models (1) and (5): $\chi^2(1)=6.39$ Prob> $\chi^2=0.0115$; models (2) and (6): $\chi^2(1)=18.96$ Prob> $\chi^2=0.0000$]



Universidad de
San Andrés

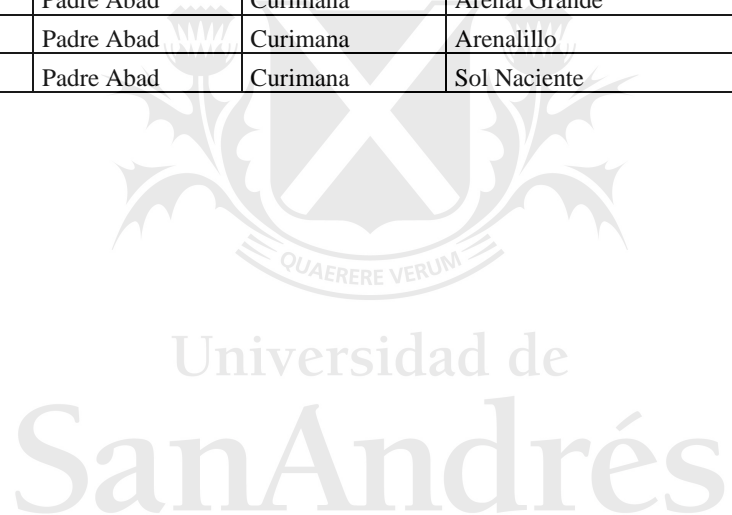
Appendix. Calculations of Topographic Distances

Department	Province	District	Population center	Topographic distance (km)
Apurimac	Abancay	Abancay	Atunpata	0.03
Apurimac	Abancay	Abancay	Quisapata	3.72
Apurimac	Abancay	Abancay	Wiracochapata	1.6
Arequipa	Caraveli	Acari	Lucasi	1.4
Arequipa	Caraveli	Acari	Santa Teresa	1.14
Arequipa	Caraveli	Atico	Chorrillos	10.85
Arequipa	Caraveli	Bella Union	San Isidro	0
Arequipa	Caraveli	Lomas	Costa Azul	0.01
Arequipa	Caraveli	Lomas	Santa Sarita	0.01
Arequipa	Caraveli	Yauca	Alto Tupac	0.01
Arequipa	Caraveli	Yauca	Yauca	0.01
Cajamarca	San Marcos	Ichocan	Illuca	6.48
Cajamarca	San Marcos	Ichocan	Llanupacha	0.1
Cajamarca	San Marcos	Ichocan	Paucamarca	3.73
Cajamarca	San Marcos	Ichocan	Paucamayo	3.58
Cajamarca	San Marcos	Ichocan	Poroporito	0.08
Cajamarca	San Marcos	Pedro Galvez	Catagon	0.4
Cajamarca	San Marcos	Pedro Galvez	Pomabamba	2.84
Cajamarca	San Marcos	Pedro Galvez	Rancho Grande	0.04
Cajamarca	San Miguel	Catilluc	Catilluc	0.52
Cajamarca	San Miguel	Catilluc	Catilluc Bajo	0.2
Cuzco	Paucartambo	Caicay	Ccollataro	0.06
Cuzco	Paucartambo	Paucartambo	Phuyucalla	0.02
Huancavelica	Tayacaja	Acraquia	Mucuro	2.3
Huancavelica	Tayacaja	Acraquia	Pamuri	0.04
Huancavelica	Tayacaja	Acraquia	San Cristobal	0.13
Huancavelica	Tayacaja	Acraquia	Tomanya	0.06
Huancavelica	Tayacaja	Salcabamba	Caymo	0.06
Huancavelica	Tayacaja	Salcabamba	Garcia Pampa	1.33
Huanuco	Huamalies	Jacas Grande	Nuevas Flores	0.2
Huanuco	Huamalies	Llata	Buena Vista	5.02
Huanuco	Huamalies	Llata	Libertad	0.46
Huanuco	Huamalies	Llata	Ocshash	0.01
Huanuco	Huamalies	Llata	Sacuatuna	1.34
Ica	Chincha	Chincha Baja	Salinas	0.02
Ica	Chincha	Chincha Baja	Valencia	0.00
Ica	Chincha	Chincha Baja	Vilma Leon	0.01
Ica	Pisco	Independencia	Cabeza De Toro Lateral 6	1.47
Ica	Pisco	Independencia	Fermin Tanguis	1.47
Ica	Pisco	Independencia	Nuevo Huanuco	1.47

Department	Province	District	Population center	Topographic distance (km)
Junin	Satipo	Mazamari	Los Angeles De Eden Alto	0.01
Junin	Satipo	Mazamari	Materiato	1.28
Junin	Satipo	Mazamari	Mirador De Cañete	0.15
Junin	Satipo	Mazamari	San Vicente De Cañete	0.10
Junin	Satipo	Rio Negro	Bajo Huahuari	0.31
Junin	Satipo	Rio Negro	Centro Hauhuari	0.16
Junin	Satipo	Rio Negro	Centro Huahuari	0.16
Junin	Satipo	Rio Negro	Santa Rosa De Panakiari	1.53
Junin	Satipo	Satipo	Alto Capiro	0.25
Loreto	M. Ramon Castilla	Caballococha	Bufeo Cocha	8.72
Loreto	M. Ramon Castilla	Caballococha	Nuevo Palestina	6.03
Loreto	M. Ramon Castilla	Yavari	Fujimori	59.31
Loreto	M. Ramon Castilla	Yavari	Rondinha Zona I	57.53
Loreto	M. Ramon Castilla	Yavari	Santa Rosa	47.15
Pasco	Oxapampa	Oxapampa	Arcuzazu	0.04
Pasco	Oxapampa	Oxapampa	El Abra	0.50
Pasco	Oxapampa	Oxapampa	Quillazu	0.40
Piura	Sullana	Lancones	El Cortezo	0.30
Piura	Sullana	Lancones	Pampas Quemadas	3.20
Piura	Sullana	Lancones	Sausal	5.00
Piura	Sullana	Sullana	Cieneguillo Norte	1.92
Piura	Sullana	Sullana	Las Lomas	1.21
Piura	Sullana	Sullana	Las Mercedes	0.04
Piura	Sullana	Sullana	San Juan De Los Ranchos	16.48
Piura	Sullana	Sullana	Santa Rosa	3.30
Piura	Sullana	Sullana	Tres Compuertas	0.04
Puno	Huancane	Cojata	Bellapampa	4.82
Puno	Huancane	Cojata	Tomapirhua	2.41
Puno	Huancane	Huancane	Bellapampa	4.82
Puno	Huancane	Huancane	Chacacruz	0.01
Puno	Huancane	Huancane	Taurahuta	0.03
Puno	Huancane	Huatasani	Catarani	6.68
Puno	Huancane	Huatasani	Ccanco	1.28
Puno	Huancane	Huatasani	Curupampa	6.10
Puno	Huancane	Huatasani	Huatapata	1.28
Puno	Huancane	Huatasani	Llinquipata	0.46
Puno	Huancane	Huatasani	Quencha Milliraya	0.05
Puno	Huancane	Huatasani	San Calvario Pongoni	1.63
Puno	Huancane	Huatasani	Tintapata	1.07
San Martin	Rioja	Nueva Cajamarca	Angaiza	0.20
San Martin	Rioja	Nueva Cajamarca	La Primavera	1.45
San Martin	Rioja	Nueva Cajamarca	Palestina	0.12

Department	Province	District	Population center	Topographic distance (km)
San Martin	Rioja	Nueva Cajamarca	Vista Alegre	0.03
San Martin	Rioja	Pardo Miguel	El Afluyente	10.19
San Martin	Rioja	Pardo Miguel	San Juan Del Mayo	2.95
Ucayali	Coronel Portillo	Yarinacocha	11 De Agosto	1.00
Ucayali	Coronel Portillo	Yarinacocha	Aahh La Capirona	0.06
Ucayali	Coronel Portillo	Yarinacocha	Aahh Monterrico	0.06
Ucayali	Coronel Portillo	Yarinacocha	Jose Olaya	0.23
Ucayali	Coronel Portillo	Yarinacocha	Las Damas De Milagro	0.06
Ucayali	Coronel Portillo	Yarinacocha	San Francisco	0.23
Ucayali	Coronel Portillo	Yarinacocha	San Jose	0.00
Ucayali	Coronel Portillo	Yarinacocha	San Juan	0.01
Ucayali	Coronel Portillo	Yarinacocha	San Lorenzo	0.30
Ucayali	Coronel Portillo	Yarinacocha	Santa Rosa	0.10
Ucayali	Padre Abad	Curimana	Arenal Grande	17.55
Ucayali	Padre Abad	Curimana	Arenalillo	17.55
Ucayali	Padre Abad	Curimana	Sol Naciente	4.00

Source: OSINERGMIN



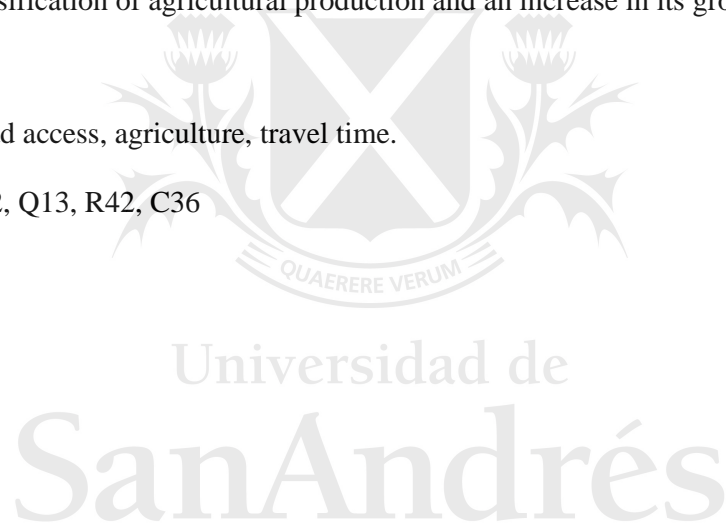
Roads and Agriculture: Impacts of Connectivity in Peru*

Julio Aguirre **Yohnny Campana** **Elmer Guerrero**
Universidad del Pacífico Universidad de San Andrés Universidad de San Andrés

Abstract: We analyze the impact of road infrastructure on agriculture using geo-referenced data on agricultural and transport sectors as well as an instrumental variables specification. We find that road connectivity reduces farmers' travel time and increases their sales. Our results are robust to the inclusion of year and region fixed effects, and also to controlling for geography (i.e. gradient and altitude), population (i.e. native language and education), and mining activity. However, we do not find statistical evidence that the provision of roads allows the diversification of agricultural production and an increase in its gross value.

Keywords: Road access, agriculture, travel time.

JEL Code: Q12, Q13, R42, C36



November 2015

* Julio Aguirre, aguirre_jc@up.edu.pe, Yohnny Campana, yohnny.campana@gmail.com, Elmer Guerrero, elmer.guerrero@gmail.com. We thank Martín Rossi, Hugo Fano, Richard Webb, Carlos Amat y León, Rosario Gómez, Daniel De La Torre Ugarte, Pablo Lavado, Héctor Maletta, and Iván Rivera for their useful comments and suggestions. This work was supported by the Food and Agriculture Organization of the United Nations (FAO), the Economic and Social Research Consortium (CIES), and the Peruvian Ministry of Agriculture and Water (MINAGRI) - Grant Number TCP/PER/3402.

I.- Introduction

Road infrastructure is extremely important in the rural areas of developing countries, where the main economic activity is agriculture. Road infrastructure has the immediate effect of reducing the time and costs of transport between production fields and marketplaces (BIDS, 2004). By reducing transportation costs, roads increase agricultural productivity, access to advanced technology, capital, and employment opportunities. Increased income leads to an increase in consumption expenditure and a reduction in poverty (Khander et al., 2006); Nguyen, 2011; van de Walle, 2009; Jacoby, 1998). Webb (2013) highlights the negative correlation between agricultural productivity and the degree of isolation of population, further aggravated by inadequate infrastructure in rural areas. In other words, less connectivity is related to the lower production development. A new transport facility that reduced the costs of transporting farm products to the end-consumer could increase the value of farms products (Schultz, 1964).

On the other hand, rural roads, as a poverty-alleviation instrument, are an important form of public policy, which involves the mobilization of financial resources not only from the government itself but also from donor institutions and multilateral agencies (van de Walle and Cratty, 2002). Given the limited budget, it is important to determine whether the dollars spent on the construction of a road has an impact that justifies the investment compared to other alternatives (e.g., an intervention in education, also believed to have considerable impact) (van de Walle, 2009).

In this paper, we analyze the impact of road infrastructure on Peru's agriculture sector. Using databases on agricultural and transport sectors, we find that the expansion of road connectivity reduces farmers' travel time (in those districts where producers on average take less than 24 hours to travel from home to district) by 3.1 hours, and increases the proportion of output sold in the districts by up to 40 percentage points -especially in larger agricultural units (farmer-holdings with an area exceeding 20 hectares). Our results are robust to the inclusion of year and region fixed effects, and also to controlling for geography (i.e. gradient and altitude), population (i.e. native language and education), and mining activity. In addition, we find that the provision of roads allows an increase in the district diversification of agricultural production index by up to 55 percentage points and an increase in its gross value, although this latter result is not entirely robust to the inclusion of other controls.

Our paper contributes to an important body of literature on rural road impacts on agriculture. Studies in Asia, Latin America and Africa find that road infrastructure has a positive impact on a number of agricultural outcomes: reduction in transport costs (Jacoby and Minten, 2009, in Madagascar), improvements in farm productivity (Kiprono and Matsumoto, 2014, in Kenya; Kingombe and di Falco, 2012, in Zambia), increases in household income and consumption (Escobal and Ponce, 2003, in Peru; Jalan and Ravallion, 2002, in China; Khander et al., 2006, in Bangladesh), crop intensification (Khander et al., 2006, in Bangladesh; van de Walle, 2009), reduction of poverty (Dercon et al., 2007, in Ethiopia; Gibson and Rozelle, 2003, in Papua New Guinea; Jalan and Ravallion, 1998, in China; van de Walle and Mu, 2011, in Vietnam) and increases in production (Dorosh et al., 2012, in Sub-Saharan Africa).

Any attempt to carry out impact assessment must address endogeneity, which arises from the selective choice of the routes of roads: if the allocation of roads in a geographic space is random, comparing the interest variables of agricultural units with access to roads to those without access is sufficient to detect impacts. However, this is not possible because the routes of roads depend on observable constraints such as the physical characteristics of the geographical areas that are to be connected, and unobservable characteristics such as strategic decisions of governments that seek to link spaces to promote their development (van de Walle, 2009; Binswanger et al., 1993; Jalan and Ravallion, 1998). Both considerations, particularly the latter, configure the endogenous character of the provision of roads and require the adoption of econometric strategies that deal with this problem in order to find consistent parameters.

We use a cross section database and an instrumental variables model, using potential routes between two cities as a source of exogeneity (different from existing roads). Faber (2014), Banerjee et al. (2012), and Martincus et al. (2013) provide an interesting frame of reference that we follow in this study.

Faber (2014) uses the Chinese National Truck Highway System (NTHS) as a large-scale natural experiment to assess whether reducing marketing costs contributes to boosting economic activity from metropolitan centers to the peripheral regions of China. The identification assumption is that the location of a district along the network of lower costs in China affects changes in economic performance at the district level only through the connections of NTHS highways, conditioned on fixed effects at the provincial level, the distance to the roadway system, and controlling for political and economic characteristics. Martincus et al. (2014) study the effects of road infrastructure on Peruvian exports and job

creation. The authors use the road system of the Incas as a source of exogenous variation of transport infrastructure (because the system was built for reasons completely unrelated to current international trade). Banerjee et al. (2012) estimate the impact of access to better infrastructure on Chinese regions. The authors use district-level economic data from China and as an identification strategy use Euclidean lines connecting historic cities in the country, reproducing the Chinese railway infrastructure built between the late nineteenth and early twentieth centuries. They identify average treated areas as those that are close to straight lines between the same set of cities. Then, the analysis compares areas near the lines to areas further away, and interprets the result of this comparison as the total effect of transport infrastructure along these historic transportation corridors.

In Peru, there are many studies linking agriculture development and road infrastructure (construction and rehabilitation), such as Webb (2013), Escobal and Torero (2005), and Escobal (2000). Their methodology is primarily correlational and the existence of road infrastructure is just one among several determinants of the agricultural variables considered. Moreover, the authors do not address the problem of endogeneity associated with the spatial placement of road infrastructure.

Escobal and Ponce (2003), though they do deal with endogeneity, only address the issue of observables through a matching estimator. However, it is important to recognize that the endogeneity of infrastructure provision must respond largely to unobservable factors, such as government decisions at different levels (central, regional, and local). Valdivia (2010) does approach this problem using a differences-in-differences estimator that addresses the problem of endogeneity in some way; however, the author uses only two years as an evaluation period after the construction of roads, which is too short a period for detecting impacts. In addition, both studies use sample information (which is also limited) from the spaces in which PROVIAS¹ carries out its activities, which are hardly representative of the rest of the country.

We consider one advantage of our paper to be that we use a more extensive definition of road provision, which is appropriate for assessing the contribution of road infrastructure to agriculture. Our explanatory variable is district road density, measured as the ratio of the total length (km) of district roads to the area (km²) of the district. This definition provides us with an approximation of the real existing network of roads that permit interdistrict, interprovincial, and interdepartmental integration. In addition, the information on road

¹ PROVIAS is a special project created by the Ministry of Transport and Communications and charged with the rehabilitation and maintenance of the national road network.

infrastructure represents a stock accumulated in the last decade, which may reveal long-term impacts.

The paper continues as follows: Section II describes the data and presents the estimation framework; Section III reports the results; and Section IV provides the conclusions.

II.- Data and estimation framework

The main source of the variables we use is the database of the Fourth National Agricultural Census carried out in 2012 (CENAGRO 2012). We construct indicators of interest from CENAGRO data to analyze the impact of road connectivity on the development and welfare of agricultural producers in the Peru. We obtain additional data on roadways from the Ministry of Transport and Communications, which has geo-referenced information on the country's road network.

Furthermore, we use data from the 1993 National Population and Housing Census and the 1994 CENAGRO to build the controls that are introduced into the regressions. More geo-referenced data is obtained from the Shuttle Radar Topography Mission (SRTM) in order to calculate the physical characteristics of the land in the district (mainly height and average grade). We obtain information on districts with mining projects from Mapa Minero 2013, available on the website of the Ministry of Energy and Mines.² In addition, we use information on per capita income, the human development index, and identification of the natural regions of Peru from the 2012 Report of Human Development (PNUD, 2013).

We collapse all databases to district level and use ArcGis software. For the treatment of road information to calculate road length within each district, we also consider the length of all roads within this geographical unit, taking into account the topography of the spaces. For this purpose, we use topographic information available on the website of SRTM NASA.³ We construct indicators of average district altitude in meters above sea level (MASL) and average gradient of the land, and measure the area in each district in square meters.

As outcome variables (A_{id}) in the econometric specifications we consider: use for production (proportion of output sold, proportion of output for self-consumption, proportion of output for own input, and proportion of output for animal feed); diversification of production; gross value of agricultural production (GVP); welfare (Human Development

² http://www.minem.gob.pe/_publicacion.php?idSector=1&idPublicacion=449

³ <http://www2.jpl.nasa.gov/srtm/>. However, spatial data used in this document are extracted directly from the statistical annex of Dell (2010) available in the shape format.

Index - HDI); logarithm of per capita income; subjective poverty (the proportion of households who consider that farming provides enough income to cover expenses); diversification of employment (the proportion of households that stopped working in their agricultural unit and seek income in other occupations); and travel time (from home to district capital and the proportion of homes in the district whose travel time to the district capital is less than 24 hours). With the exception of the HDI and per capita income variables (which are obtained from the 2012 Report of Human Development) and the GVP (which is obtained from the Ministry of Agriculture and Irrigation - MINAGRI), the source for the remaining outcome variables is CENAGRO 2012.

We use other variables as regressors. The first is road density, measured as the length of roads within each district divided by the area of the district. In addition, in order to isolate the influence of factors other than road networks, we use information about the existence of mining projects in a district, and data on average altitude and inclination of the terrain, among others. In Table 1 we report the descriptive statistics of the variables.

Insert Table 1

Identification strategy

In this study, we use an empirical strategy inspired by those used by Faber (2013), Banerjee et al. (2012), and Martincus et al. (2012). These researchers employ an estimator of instrumental variables (IV) exploited as a source of exogenous Euclidean straight lines that would connect the spaces if the terrain of the land was completely flat. In particular, the strategy of these authors is based on the idea that the objective of road infrastructure is to connect major cities, i.e. high nodes of population concentration or places of importance for other reasons. Thus, if the topography is flat and there are no restrictions on any desirable part of the route, then the shortest path of a road (or track) should follow a straight Euclidean line between each pair of points so that the path to connect each city is minimal, yet efficient in the sense that it involves constructing a limited number of pathways. However, this is not possible, because the effective route is conditioned by the characteristics of the terrain, soil conditions, and other factors that influence the feasibility of building the infrastructure. But ultimately, a Euclidean straight line should be a benchmark regarding the construction of an effective route. Consequently, this reference line may be used as a tool for modeling the effective route.

Using ArcGis software, we estimate the minimum distance between the centroid of the geographic units of analysis (districts) and the closest point to the reference line (route).

Map 1 shows a tree of routes that connect the capitals of all departments of Peru (black lines). National roads are denoted in blue, departmental roads in red, and local roads in green. Clearly, the graph appears to show that the highest concentration of routes occurs along straight lines. We will subsequently test this empirically.

Insert Map 1

Then, in the estimates, we use this instrument to estimate the following first stage of estimator IV:

$$V_{id} = \alpha + \beta \ln R_{id} + X_{id}\eta + \omega_r + e_{id} \quad (1)$$

Where V_{id} is the road density in the geographic unit of analysis (district) i of department d ; $\ln R_{id}$ is logarithm of the distance between the centroid of district i and the closest point of the tree; X_{id} is the set of controls that vary at the level of the geographical unit i or the department; ω_r are fixed effects at the level of the natural region (eight regions proposed by Javier Pulgar Vidal in 1938); and e_{id} is the error term. Subsequently, the second stage of the estimator is:

$$A_{id} = \gamma + \beta V_{id} + X_{id}\phi + \omega_r + u_{id} \quad (2)$$

Where A_{id} is the agricultural outcome variable of interest and β is the parameter of interest, which captures the causal effect of road connectivity. We compute standard errors clustered at the department level.

Given the methodology, in order for parameter β to be consistent in Equation 2, it is necessary for two assumptions to hold (Wooldridge 2010). The first is the assumption of relevance, which requires that $\ln R_{id}$ be strongly correlated with V_{id} . That is, the proximity of a district to the straight lines in the Euclidean tree allow us to predict existing road density. This assumption can be tested empirically using Equation 1 above. The second assumption, called exclusion, requires that the lines on the Euclidean tree not be directly correlated with the residue of Equation 2. That is, the only channel through which this instrument should affect the agricultural outcome variables is through the provision of road infrastructure.

This second assumption deals with identification, so it cannot be tested directly. However, it could be violated if the districts located along the lines of the Euclidean tree consistently had characteristics that were more favorable for agricultural and/or economic development. Therefore, to control for this possibility, in the set of regressors X_{id} characteristics of geographical units are introduced, such as the gradient and average altitude of the land as well as characteristics that measure the endowment of human capital and the percentage of people with different educational levels. In addition, given that the districts that

are closer to city nodes are automatically more connected but also the most prosperous, we exclude from the analysis those that fall within a radius of 50 km of these points⁴ (in other words, in these spaces monotone road density is higher than in other districts because these are large urban cities. Map 2 displays in black those districts that were left out of the estimate sample for this reason. These constitute a total of 640 districts, and thus the maximum usable sample is 1,194 districts

Insert Map 2

III.- Results

3.1.- First stage⁵

The relevance assumption of the instrumental variables estimator requires that the instrument be correlated with the endogenous regressor. In this particular case, this means that the distance between the district and the line nearest the Euclidian tree should be a good predictor of road density at the district level. Graph 1 confirms this characteristic. It can be seen that the more distant districts in the Euclidian straight projection show a lower road density, while those that are closer have more km of roads per km². This is also confirmed in Table 3. After estimating equation (1) the coefficient δ is statistically significant at 1% and has a magnitude of about -0.05 and robust to the inclusion of geographical, population or mining controls. Moreover, the F statistic is above 10, indicating that the model does not have a problem of weak instrument bias.

Insert Graph 1

Insert Table 3

3.2.- Second stage

In tables 4 to 6 and 9 we report estimates of Equation (2), and in table 8 we present an additional examination to test the robustness of results. In addition, given that the large number of measured outcomes raises concerns about multiple inference, significant coefficients may emerge simply by chance, even if there are no treatment effects. This problem is well known in the theoretical literature (Romano and Wolf 2005) and the biostatistics field (Hochberg 1988). For inference, we additionally perform joint tests of the hypotheses following the Anderson (2007) Multiple Testing Procedure.

⁴ Faber (2013) adopts the same strategy.

⁵ See in Appendix some insight about the correlation of the instrument.

Uses for production

In Table 4 we report estimates of equation (2) for the outcome variables of production uses: the proportion of cultivated hectares intended primarily for sale, consumption, own-input or feed for animals. Thus, the estimated coefficients measure the rate at which production would change as a result of a marginal increase in road density.

Insert Table 4

In columns (1) and (6) of Table 4, we present the OLS estimations for the whole country and the Andes, respectively. Clearly, OLS coefficients underestimate the impact of road connections on agricultural sales, but overestimate the impact on other variables. Theoretically, the positive bias of OLS could be explained because the unobservable variables are negatively correlated with the provision of roads and positively correlated with market sales.

In column (2) of Table 4, we report the IV estimations. These regressions are controlled only for the age of district and the fixed effects of the natural regions defined by Javier Pulgar Vidal. The coefficients indicate that the provision of roads allows districts to have a higher quantity of sales and reduce the area designated for alternative uses. Thus, apparently, the provision of roads has adequately articulated agricultural production with markets. These results are in line with Webb's (2013) assertions. However, in order to control for a possible influence of the physical aspects of the land that affect the results, we added to the regressions some physical characteristics of the land such as the gradient of the terrain and average altitude, so as to capture the productive potential of the land. The results reported in column (3), respectively, remain little changed.

In addition, initial differences in the characteristics of the population could explain agricultural potential and/or greater willingness to enter the market. To control for this, we added regressors that capture the percentage of rural population in 1993, the percentage of population with a native language other than Spanish (Quechua, Aymara, or other) in 1993, the percentage of population with secondary or higher education in 1993, the percentage of population engaged in agriculture in 1993, and the logarithm of the population in 1993 (to control for the initial allocation of the labor force). Column (4) present the parameters with slight variations, but the magnitude and direction remain similar.

In recent times, mining has become an activity with considerable influence on agricultural activity (Zegarra et al., 2007). We use a dummy variable as a control that takes

the value of 1 for those districts that had mining activity in 2014. Results in column (5) demonstrate that the inclusion of this variable does not greatly change the results.

Finally, in Table 4 we present the results of the other production use outcome variables. They clearly show that the largest increase in the proportion of land employed for products to sell on the market comes primarily at the cost of reduced consumption. Additionally, a reduction in own-input is observed. No impacts are seen in relation to animal feed.

Gross value of production (GVP) in 2009 and diversification of production

Table 5 reports the estimations for the logarithm of the GVP. According to the results, the impact of road infrastructure on that outcome variable is positive for OLS and IV specifications – column (1) and (2), respectively. However, the parameter is not robust to the inclusion of controls. Thus, the evidence is inconclusive regarding the effect of the provision of roads on agricultural production.

Insert Table 5

On the other hand, we consider an additional variable indicative of the diversification of production, measured as the entropy index $E = -\sum_{j=1}^n w_j \cdot \ln(w_j)$, where w_j is the share of product j in the total cultivated area (Foster and Jara, 2005). The impacts of this variable are not auspicious than the previous one since the parameters exhibit high variability to the inclusion of variables.

Welfare

Table 6 presents the estimations for proxy variables for welfare outcome variables. The parameters indicate that impacts are statistically null. Thus, the parameters are similar to those found for production uses (mainly sales).

Insert Table 6

Additional examinations:

Heterogeneous effects

It is reasonable to expect that the impacts have also operated heterogeneously from dimensions that reflect the ability of farmers to seize the opportunities that come with the provision of roads. Economies of scale can occur because the average cost per unit of production decreases as the size of the farm increases, given farmers are able to spread more production over the same level of fixed expenses. Therefore, we made an effort to detect heterogeneous impacts depending on the scale of operations of agricultural units, measured

by the size of the land under exploitation. Thus, we classify the population into three groups: small farmers – landholdings of less than 5 hectares under exploitation; medium farmers – holdings of between 5 and 20 hectares; and large farmers – holdings with an area exceeding 20 hectares under exploitation. Then we calculate the distribution of land cultivated by production uses (sales, own-input, auto-consumption, and animal feed) for each district within the three categories of producers. The results are shown in Table 8. The impacts appear to have affected the three groups of producers but the larger scale agricultural units report the highest effects.

Insert Table 8

Travel time

Reducing the travel time between the point of origin and a destination is the main immediate effect generated by the provision of and access to roads. Table 9 reports the effect of the increase in road density in districts in two outcome variables: (i) the travel time of the agricultural producer from home to the capital of district⁶; and (ii) the proportion of households that need less than 24 hours to journey from home to the district capital. As can be seen, in the OLS and IV models the coefficients have the expected signs and are statistically significant. That is, these coefficients indicate that the provision of roads leads to a reduction in travel time of agricultural producers from home to the district capital, and also reduces the percentage of households in the district that need over 24 hours to move between these same points.

Insert Table 9

IV. Conclusions

In general, the provision of roads (including maintenance and/or improvement) may create opportunities for economic growth and poverty reduction through various mechanisms: reduced transportation costs, improved access to markets and technology, expansion of agricultural and non-agricultural production, and greater access to inputs. Also, on the agricultural household level, the development of road infrastructure contributes to increased productivity, income, health levels, and access to education.

This study evaluates the impact of the provision of roads in Peru's Agriculture sector. In recent years there has been an explosion in the construction of roads around the country

⁶ The administrative divisions of Peru, from smallest to largest, until recently were districts, provinces, and departments. Departments were replaced by regions in 2002 but Peruvian ministries still sometimes refer to departments when gathering information.

which has tended to connect the most isolated markets to big city spaces. These connections have led to a transformation of the rural areas of the country (Webb, 2013).

Our estimations report the following results:

- a) The distance between the district and the nearest straight Euclidian tree constitutes a good predictor of road density at the district level - in the districts that are more distant from the straight Euclidean lines connecting departmental capitals, there is lower road density, while in those that are closest, there are more kilometers of roads per km².
- b) More road connectivity increases sales and reduces self-consumption. In fact, according to the estimations, for every kilometer of roads built per square kilometer of surface of the district, market-oriented production increases up to 36 percentage points in the country.
- c) Road connectivity has allowed for the diversification of agricultural production and an increase in the gross value of agricultural production, although these results are not entirely robust to the inclusion of potentially explicative factors other than roads. Moreover, there is only weak evidence that all this had led to higher per capita household incomes. These results are similar to those in Faber (2013), who argues that this could be evidence of the reallocation of resources that occurs as a result of road construction, which allows for the flow of factors to places where these have a higher return;
- d) In terms of welfare, the results are statistically invalid.
- e) The greatest impacts are concentrated among larger agricultural units. In the case of small units, even when the impacts are positive, they do not reach the level found for large units.
- f) For those districts in which producer travel time from home to the district capital is less than 24 hours, the farmer's travel time reduces by up to 3.1 hours if the average road density district is increased by 1 Km/Km². In addition, the proportion of households that require less than 24 hours to travel from home to the district capital is reduced by up to 10 percentage points.

Our results are very eloquent. On the one hand, clear impacts can be seen on production uses (market sales), but this do not necessarily correlate with an increase in the production sector. This can be explained by the fact that agricultural development not only requires provision of roads, but also a package of infrastructure intended to provide the

necessary conditions so that producers can offer quality products (greater supply and timely maintenance of irrigation infrastructure, increased micro reservoirs, transparent management by organizations of water users, among others). This is in line with studies by Escobal (1995) and Valdivia (2010), according to which roads are a key but insufficient factor for the development of spaces. Rather, complementary public goods are needed to enhance the effects of roads on the welfare of the population, beyond enabling them to increase their access to markets.

On the other hand, following the interpretation of Faber (2013), our results show that despite the provision of roads, the impact on investment in capital assets or production technology is very limited, which could be because the benefits of major roads and better connectivity tend to migrate, in part, to the larger cities, where the yield is higher. In that sense, there is room to develop measures to encourage investment in infrastructure to enhance the impact of roads.

Finally, greater provision of roads (and their improvement) is not all that is necessary and is not even the most important task in the long term. Following Iguñiz (1998), those initiatives that will increase the movement of products with high value per unit of weight⁷ - whether they are chosen for characteristics that they already possess or whether they are transformed to increase their value - should be developed. The lower the cost of transport, the lower the competitive advantage needed by producers to stay close to final markets. Certainly, such initiatives could be promoted by the Ministry of Agriculture and Irrigation (MINAGRI) or the Ministry of Production (PRODUCE).

References

AFIN (2012). Por un Perú Integrado. Plan Nacional de Infraestructura 2012-021”, Asociación para el Fomento de la Infraestructura Nacional, Instituto de Regulación de la Universidad ESAN and Universidad del Pacífico, Lima.

Anderson, M. (2008). “Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects”, *Journal of the American Statistical Association*. 103(484): pp. 1481-1495.

Banerjee, A. Duflo, E. and Qian, N. (2012). “On the Road: Access to transportation infrastructure and economic growth in China.” NBER Working Paper 17897, March.

⁷ Every time a road is constructed – and paved – a large reduction in time and costs of transport is achieved; after this, however, further reductions in costs of transport are slow and marginal.

BIDS (2004). “Poverty Impact of Rural Roads and Markets Improvement and Maintenance Project of Bangladesh”, Bangladesh Institute of Development Studies *mimeo*.

Binswanger, H.P., Khander, S.R. and Rosenzweig, M. (1993). “How Infrastructure and Financial Institutions Affect Agricultural Output and Investment in India,” *Journal of Development Economics* 41: 337-366.

Chauvin, N.D. and Porto, G. (2011). “Market Competition in Export Cash Crops and Farm Income,” *African Center for Economic Transformation and Universidad de La Plata*.

Deininger, K. and Okidi, J. (2003). “Growth and Poverty Reduction in Uganda, 1999 – 2000: Panel Data Evidence”, *Development Policy Review*, Vol. 21, pp. 481-509.

Dercon, S., Gilligan, D. O., Hoddinott, J., & Woldehanna, T. (2007). The impact of roads and agricultural extension on consumption growth and poverty in fifteen Ethiopian villages [CSAE Working Paper Series] (Tech. Rep. Nos. 2007–01). Centre for the Study of African Economies, University of Oxford.

Escobal, Javier, and Maximo Torero (2005). ‘Adverse Geography and Differences in Welfare in Peru’, in R. Kanbur and A. J.Venables (eds), *Spatial Inequality and Development*, Oxford University Press: Oxford

Escobal, J., & Ponce, C. (2003). The benefits of rural roads. Enhancing income opportunities for the rural Poor [Working Paper] (Tech. Rep.). Grupo de Análisis para el Desarrollo (GRADE)

Faber, B. (2014). “Trade Integration, Market Size, and Industrialization: Evidence from China’s National Trunk Highway System”, *Review of Economic Studies* (2014) 81, 1046–1070.

Fan, S. and C. Chan-Kang (2004). “Road Development, Economic Growth, and Poverty Reduction in China,” *IFPRI DSGD Discussion Paper* 12, 86.

Fan, S., Hazell, P. and Thorat, S. (2000). “Government Spending, Growth, and Poverty in Rural India,” *American Journal of Agricultural Economics*, 82(4): 1038-1051.

Foster, W. and Jara, E. (2005). “Diversificación de exportaciones agrícolas en América Latina y El Caribe: patrones y determinantes,” *Economía Agraria*, Vol. 9, Pontificia Universidad Católica de Chile, Santiago.

Fosu, K. N. (1995). *Public Goods and Services and Food Security: Theory and Modeling Approaches with Special Reference to Ghana and Burkina Faso*. Réseau SADO Seminar. Accra Ghana.

Gannon, C. and Liu, Z. (1997). “Notes for the Impact Evaluation of the Construction and Rehabilitation of Rural Roads”, *mimeo*, DIME, World Bank, Washington, D.C.

Hochberg, Y. (1988), “A Sharper Bonferroni Procedure for Multiple Tests of Significance,” *Biometrika*, 75, 800–802.

Iguñiz, J. (1998). "Aplanar los Andes y otras propuestas", Instituto Bartolomé de las Casas – Rímac, Centro de Estudios y Publicaciones, Febrero.

Jacoby, H. G., 1998. "Access to markets and the benefits of rural roads", Policy Research Working Paper Series 2028, The World Bank, Washington D.C.

Jacoby, H. G. and Minten, B. (2009). "On Measuring the Benefits of Lower Transport Costs," *Journal of Development Economics*, 89, 28-38.

Jalan, J., and Ravallion, M. (2002). Geographic poverty traps? A micro model of consumption growth in rural China. *Journal of Applied Econometrics*, 17 , 329–346.

Jalan, J. and Ravallion, M. (1998). "Are There Dynamic Gains from a Poor-area Development Program?", *Journal of Public Economics* 67, 65 – 86.

Khander, S.R., Bakht, Z. and Koolwal, G.B. (2006). "The Poverty Impact of Rural Roads: Evidence from Bangladesh," World Bank Policy Research Paper 3875, April, Washington D.C.

Kingombe, C.K.M. and di Falco, S. (2012). "The Impact of a Feeder Road Project on Cash Crop Production in Zambia's Eastern Province between 1997 and 2002," Graduate Institute of International and Development Studies Working Paper N° 04, Génova.

Levy, H. (1996). "Kingdom of Morocco: Impact Evaluation Report, Socioeconomic Influence of Rural Roads," Operation Valuation Department, World Bank: Washington DC.

Martincus, C.V., Carballo, J. and Cusolito, A. (2014). "Routes, Exports, and Employment in Developing Countries: Following the Trace of the Inca Roads", Inter-American Development Bank, mimeo.

Mu, R., & van de Walle, D. (2011). Rural roads and local market development in Vietnam. *The Journal of Development Studies*, 47 (5), 709–734.

Nguyen, C.V. (2011). "Estimation of the impact of rural roads on household welfare in Viet Nam", *Asia-Pacific Development Journal*, Vol. 18, N° 2, December.

Qin, Y. and Zhan, X. (2012). "The Road to Specialization in Agricultural Production. Evidence from Rural China", Development Strategy and Governance Division, IFPRI Discussion Paper 01221, October.

Romano, J., and Wolf, M. (2005), "Stepwise Multiple Testing as Formalized Data Snooping," *Econometrica*, 73, 1237–1282.

Schultz, T. W. (1964). "Transforming Traditional Agriculture", New Haven and London Yale University Press.

Valdivia, M. (2010). "Contracting the road to development: Early impacts of a rural roads program, poverty and economic policy research network," PMMA Working Paper 2010-18.

van de Walle, D. and Cratty, D. (2002), "Impact Evaluation of a Rural Road Rehabilitation Project", World Bank, January, *mimeo*.

Van de Walle, D. (2009). "Impact evaluation of rural road projects," *Journal of Development Effectiveness*, 1, 15-36.

Van De Walle, D. (1996). *Infrastructure and Poverty in Viet Nam*. Washington D.C.: World Bank.

Wharton, C. (1967): "The Infrastructure for Agricultural Growth", in *Agricultural Development and Economic Growth*, ed. by H. M. Southworth, and B. F. Johnston. Ithaca, New York: Cornell University Press.

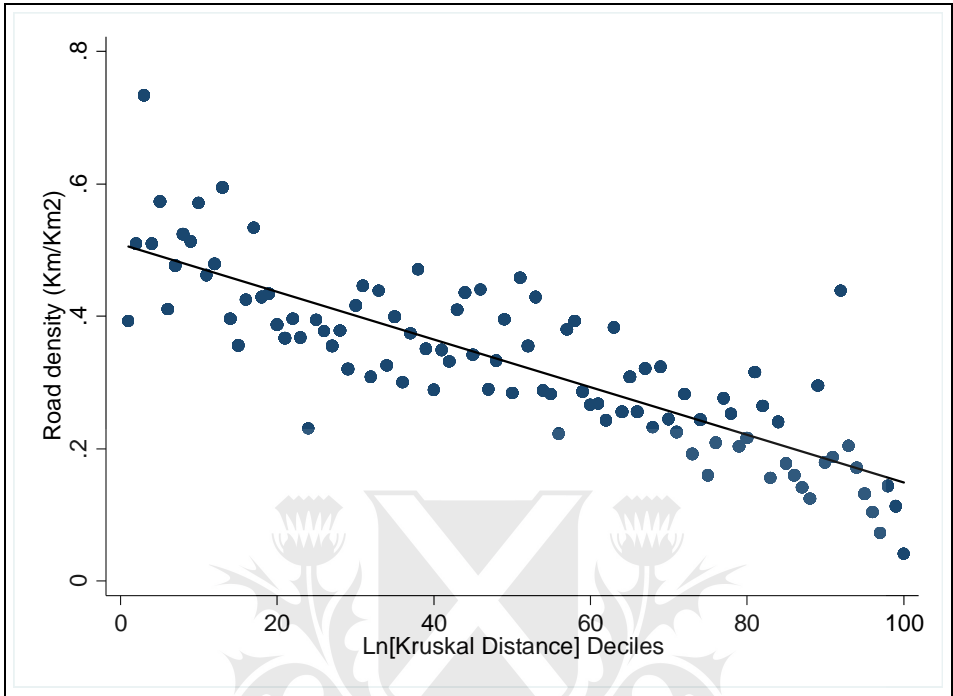
Webb, R. (2013). "Conexión y Despegue Rural," Instituto del Perú, Universidad de San Martín de Porres, Lima.

Zegarra, E.; Orihuela, J.C.; Paredes, M. (2007). "Minería y economía de los hogares en la sierra peruana: impactos y espacios de conflicto", Lima: GRADE. 85 p. Documento de trabajo, 51.



Universidad de
San Andrés

Graph 1: Instrument versus Density of Roads



Universidad de
San Andrés

TABLE 1
DESCRIPTIVE STATISTICAL SUMMARY OF VARIABLES USED

Description	OBS	Coast		Andes		Amazonian Lowlands	
		Mean	SD	Mean	SD	Mean	SD
Average ha in use by producers in district	1,784	19.4	56.0	32.3	79.5	33.1	90.3
Proportion of men in district	1,783	0.7	0.1	0.7	0.1	0.8	0.1
Average age of producer	1,783	55.3	3.8	52.3	4.3	45.5	4.1
Proportion with no schooling	1,783	0.1	0.1	0.1	0.1	0.1	0.1
Proportion with preschool	1,783	0.0	0.0	0.0	0.0	0.0	0.0
Proportion with primary school incomplete	1,783	0.2	0.1	0.3	0.1	0.3	0.1
Proportion with primary school completed	1,783	0.2	0.1	0.2	0.1	0.3	0.1
Proportion with HS incomplete	1,783	0.1	0.1	0.1	0.0	0.1	0.1
Proportion with HS completed	1,783	0.2	0.1	0.2	0.1	0.1	0.1
Proportion with non-university higher education incomplete	1,783	0.0	0.0	0.0	0.0	0.0	0.0
Proportion with non-university higher education completed	1,783	0.1	0.1	0.0	0.0	0.0	0.0
Proportion with college education incomplete	1,783	0.0	0.0	0.0	0.0	0.0	0.0
Proportion with college education completed	1,783	0.1	0.1	0.0	0.0	0.0	0.0
Proportion with Quechua as native language	1,783	0.1	0.1	0.4	0.4	0.1	0.3
Proportion with Aymara as native language	1,783	0.0	0.1	0.0	0.2	0.0	0.0
Proportion with Ashanika as native language	1,783	0.0	0.0	0.0	0.0	0.0	0.1
Proportion with another native language (not Spanish)	1,783	0.0	0.0	0.0	0.0	0.1	0.2
Proportion with Spanish as native language	1,783	0.9	0.1	0.5	0.4	0.8	0.3
Proportion with a foreign native language	1,783	0.0	0.0	0.0	0.0	0.0	0.0
Average number of crops per agricultural unit in district	1,784	3.3	2.3	4.4	1.8	3.2	0.7
Average number of plots in each agricultural unit in district	1,784	2.3	1.9	3.7	2.4	1.8	0.7
Hectares under cultivation	1,784	3,160	4,506	1,279	2,418	5,148	6,993
Sales: total crops in district	1,784	1,295	1,618	647	1,091	2,326	2,884
Self-consumption: Total crops in district	1,784	140	252	1,583	2,795	1,068	1,363
Own-inputs: Total crops in districts	1,784	25	79	35	208	39	78
Animal feed: Total crops in district	1,784	106	262	409	943	325	410
Mining exists in district	1,784	0.1	0.3	0.1	0.3	0	0.2
Estimated population in 2012	1,779	45,009	94,954	7,355	17,191	14,225	20,457
Human Development Index 2012	1,779	0.5	0.1	0.3	0.1	0.3	0.1
Life expectancy at birth	1,779	76.8	3.3	71.4	5.5	73	3.9
Population with HS completed	1,779	64	15	45	22	31	16
Years of education (population over 25 years of age)	1,779	9	2	6	2	6	1
Per capital household income	1,779	644	228	313	202	380	209

Asphalted roads in km	1,768	17,632	21,202	3,615	10,357	3,725	10,706
Unpaved roads in km	1,768	4,758	11,573	7,290	14,113	6,737	14,942
Vehicular roads in km	1,768	21,903	29,199	15,945	20,392	13,659	24,003
Roads under construction in km	1,768	0.0	0.0	37.1	655.1	0.0	0.0
Average area of districts in km ²	1,768	461	677	289	343	2,245	4,526
Average altitude of districts in meters	1,766	541	538	3,560	673	1,388	994
Average gradient in districts in degrees	1,766	3.7	3	7.8	3.5	6.1	4.4

Sources: CENAGRO 2012; MTC; NASA.



Universidad de
San Andrés

TABLE 3
PERU: FIRST STAGE ESTIMATION RESULTS (OLS)

	(1)	(2)	(3)	(4)
Ln[Euclidian]	-0.0592*** (0.0073)	-0.0578*** (0.0074)	-0.0538*** (0.0076)	-0.0529*** (0.0076)
Obs	1,175	1,153	1,102	1,102
R ²	0.117	0.132	0.158	0.160
F-stat	20.48	16.34	15.67	15.43
Geographic Variables	No	Yes	Yes	Yes
Population variables	No	No	Yes	Yes
Mining	No	No	No	Yes

Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Note: Standard errors clustered at the department level. All regressions include as regressors the age of the district (from the year of its establishment) and fixed effects for the natural regions of Pulgar Vidal. The geographic variables are the average gradient of the land and the average altitude. In addition, a logarithm of the agricultural area in 1994 has been included. The population variables include the percentage of rural population in 1993, the percentage of the population whose native language was Quechua, Aymara or an Amazonian lowland language in 1993, the percentage of the population with a high school or higher education in 1993, the percentage of the population working in the agricultural sector in 1993, and the logarithm of the population in 1993. Mining is a dummy variable which indicates if the district has mining activity.

TABLE 4
PERU: PRODUCTION DESTINATION

Variables	(1) OLS	(2) IV	(3) IV	(4) IV	(5) IV
Sales	0.163** (0.031)	0.362** (0.135)	0.322** (0.139)	0.379** (0.157)	0.381** (0.160)
R ²	0.501	0.482	0.493	0.512	0.511
Adjusted p-value	0.001	0.002	0.015	0.015	0.017
Self - Consumption	0.019 (0.031)	-0.271** (0.138)	-0.334** (0.144)	-0.414** (0.160)	-0.424** (0.164)
R ²	0.315	0.264	0.248	0.289	0.285
Adjusted p-value	0.127	0.010	0.012	0.005	0.006
Own-Input	-0.010 (0.005)	-0.056** (0.027)	-0.066** (0.029)	-0.078** (0.035)	-0.077** (0.036)
R ²	0.030	-	-	-	-
Adjusted p-value	0.198	0.117	0.227	0.331	0.352
Feed for animals	-0.172*** (0.022)	-0.034 (0.131)	0.078 (0.137)	0.112 (0.146)	0.120 (0.149)
R ²	0.213	0.195	0.194	0.228	0.225
Adjusted p-value	0.001	0.279	0.166	0.185	0.199
Obs	1,155	1,155	1,140	1,090	1,090
Geographic variables	No	No	Yes	Yes	Yes
Population variables	No	No	No	Yes	Yes
Mining	No	No	No	No	Yes

Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Adjusted p-values using Anderson (2007) Multiple Testing Procedure (See in Appendix, Table 11, original p-values). Standard errors clustered at the department level. The parameters of each variable are, respectively, the estimated coefficient/standard error of the estimated parameter and the coefficient of determination of regression R². Note: All the regressions include as regressors the age of the district in years (since its establishments) and the fixed effects for the natural regions of Pulgar Vidal. The geographic variables are the average gradient of the land and the average altitude. In addition, a logarithm of the agricultural area in 1994 has been included. The population variables include the percentage of rural population in 1993, the percentage of the population whose native language was Quechua, Aymara or an Amazonian lowland language in 1993, the percentage of the population with a high school or higher education in 1993, the percentage of the population working in the agricultural sector in 1993, and the logarithm of the population in 1993. Mining is a dummy variable which indicates if the district has mining activity.

TABLE 5
PERU: PRODUCTION (IMPACT ON THE GROSS VALUE OF
AGRICULTURAL PRODUCTION AND DIVERSIFICATION)

Variables	(1) OLS	(2) IV	(3) IV	(4) IV	(5) IV
Diversification index	0.114** (0.062)	0.891** (0.294)	0.435 (0.275)	0.400 (0.281)	0.392 (0.288)
R ²	0.068	-	0.149	0.199	0.200
Adjusted p-value	0.034	0.005	0.129	0.171	0.209
Obs.	1,155	1,155	1,140	1,090	1,090
Ln[Gross Value of Agricultural production 2009]	0.370** (0.152)	1.998** (0.702)	-0.141 (0.490)	-0.429 (0.556)	-0.548 (0.564)
R ²	0.251	0.183	0.633	0.635	0.635
Adjusted p-value	0.008	0.004	0.310	0.259	0.203
Obs.	1,151	1,151	1,139	1,088	1,088
Geographic variables	No	No	Yes	Yes	Yes
Population variables	No	No	No	Yes	Yes
Education and Mining	No	No	No	No	Yes

Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Adjusted p-values using Anderson (2007) Multiple Testing Procedure (See Appendix, Table 12, original p-values). Standard errors clustered at the department level. The parameters of each variable are, respectively, the estimated coefficient/standard error of the estimated parameter and the coefficient of determination of regression R². Note: All the regressions include as regressors the age of the district in years (since its establishment) and the fixed effects for the natural regions of Pulgar Vidal. The geographic variables are the average gradient of the land and the average altitude. In addition, a logarithm of the agricultural area in 1994 has been included. The population variables include the percentage of rural population in 1993, the percentage of the population whose native language was Quechua, Aymara or an Amazonian lowland language in 1993, the percentage of the population with a high school or higher education in 1993, the percentage of the population working in the agricultural sector in 1993, and the logarithm of the population in 1993. Mining is a dummy variable which indicates if the district has mining activity.

Universidad de
San Andrés

TABLE 6
PERU: IMPACT ON WELFARE VARIABLES

VARIABLES dependientes	(1) OLS	(2) IV	(3) IV	(4) IV	(5) IV
Human Development Index	0.006 (0.010)	-0.079 (0.050)	-0.019 (0.050)	-0.062 (0.049)	-0.042 (0.049)
Adjusted p-value	0.323	0.060	0.226	0.128	0.210
Obs.	1,175	1,175	1,153	1,102	1,102
Ln[Per capita income]	0.019 (0.055)	-0.023 (0.253)	0.232 (0.263)	0.125 (0.269)	0.249 (0.269)
Adjusted p-value	0.266	0.340	0.135	0.337	0.249
Obs.	1,175	1,175	1,153	1,102	1,102
Sufficiency of agricultural income	0.025 (0.019)	0.083 (0.067)	0.012 (0.068)	0.025 (0.079)	0.024 (0.080)
Adjusted p-value	0.254	0.336	0.998	1.000	1.000
Obs.	1,167	1,167	1,149	1,098	1,098
Diversification of employment	-0.010 (0.024)	-0.095 (0.094)	-0.004 (0.098)	-0.000 (0.107)	0.008 (0.109)
Adjusted p-value	1.000	0.574	1.000	1.000	1.000
Obs.	1,167	1,167	1,149	1,098	1,098
Geographic variables	No	No	Yes	Yes	Yes
Population variables	No	No	No	Yes	Yes
Education and Mining	No	No	No	No	Yes

Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Adjusted p-values using Anderson (2007) Multiple Testing Procedure (See in Appendix, Table 13, original p-values). Standard errors clustered at the department level. For each variable, the parameters are, respectively, the estimated coefficient/standard error of the estimated parameter and the number of observations in each regression. Note: All the regressions include as regressors the age of the district in years (since its establishment) and the fixed effects for the natural regions of Pulgar Vidal. The geographic variables are the average gradient of the land and the average altitude. In addition, a logarithm of the agricultural area in 1994 has been included. The population variables include the percentage of rural population in 1993, the percentage of the population whose native language was Quechua, Aymara or an Amazonian lowland language in 1993, the percentage of the population with a high school or higher education in 1993, the percentage of the population working in the agricultural sector in 1993, and the logarithm of the population in 1993. Mining is a dummy variable which indicates if the district has mining activity.

TABLE 8
PERU: PRODUCTION DESTINATION: HETEROGENOUS IMPACTS

Dependent variables	Small Producers			Medium-Sized Producers			Large producers		
	(1) OLS	(2) IV	(3) IV	(4) OLS	(5) IV	(6) IV	(7) OLS	(8) IV	(9) IV
Sales	0.149*** (0.031)	0.330*** (0.136)	0.326** (0.162)	0.199*** (0.031)	0.439*** (0.141)	0.440** (0.171)	0.231*** (0.041)	0.556*** (0.161)	0.560*** (0.191)
Adjusted p-value	0.001	0.007	0.042	0.001	0.001	0.011	0.001	0.001	0.009
Self-consumption	-0.026 (0.033)	-0.312** (0.145)	-0.479*** (0.176)	-0.032 (0.030)	-0.301** (0.140)	-0.459*** (0.167)	-0.013 (0.037)	-0.277** (0.135)	-0.386** (0.159)
Adjusted p-value	0.165	0.015	0.005	0.092	0.011	0.005	0.366	0.026	0.019
Own-input	-0.006 (0.003)	-0.026 (0.015)	-0.033 (0.019)	-0.005 (0.003)	-0.012 (0.015)	-0.022 (0.017)	-0.002 (0.011)	-0.059 (0.044)	-0.082 (0.055)
Adjusted p-value	0.147	0.168	0.182	0.254	0.667	0.458	1.000	0.367	0.367
Feed for animals	-0.117*** (0.020)	0.008 (0.110)	0.186 (0.134)	-0.163*** (0.023)	-0.126 (0.122)	0.041 (0.141)	-0.216*** (0.030)	-0.220 (0.156)	-0.092 (0.175)
Adjusted p-value	0.001	0.208	0.241	0.001	0.065	0.388	0.001	0.129	0.430
Obs	1,144	1,144	1,082	1,135	1,135	1,074	1,059	1,059	1,004
Geographic variables	No	No	Yes	No	No	Yes	No	No	Yes
Population variables	No	No	Yes	No	No	Yes	No	No	Yes
Mining	No	No	Yes	No	No	Yes	No	No	Yes

Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Adjusted p-values using Anderson (2007) Multiple Testing Procedure (See in Appendix, Table 15, original p-values). Standard errors clustered at the department level. For each variable, the parameters are, respectively, the estimated coefficient and the associated standard error. Note: All the regressions include as regressors the age of the district in years (since its establishment) and the fixed effects for the natural regions of Pulgar Vidal. The geographic variables are the average gradient of the land and the average altitude. In addition, a logarithm of the agricultural area in 1994 has been included. The population variables include the percentage of rural population in 1993, the percentage of the population whose native language was Quechua, Aymara or an Amazonian lowland language in 1993, the percentage of the population with a high school or higher education in 1993, the percentage of the population working in the agricultural sector in 1993, and the logarithm of the population in 1993. Mining is a dummy variable which indicates if the district has mining activity.

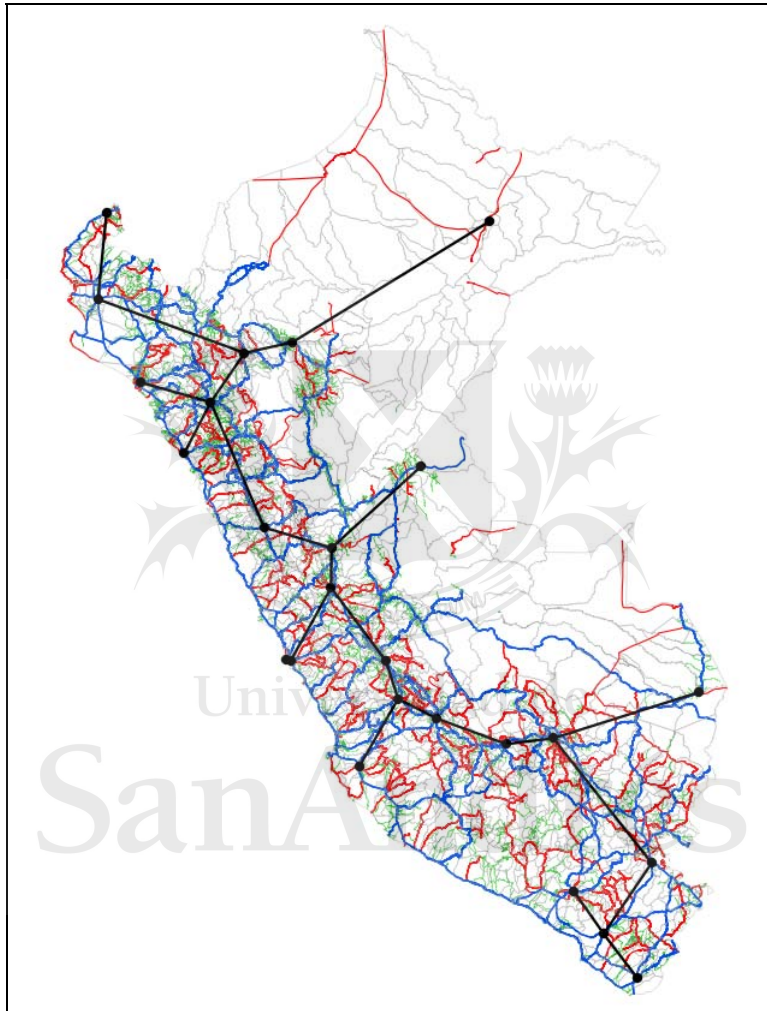
TABLE 9**PERU: TRAVEL TIME**

Dependent Variable	(1) OLS	(2) IV	(3) IV	(4) IV	(5) IV
Travel time from home to district capital	-1.740*** (0.147)	-2.225*** (0.591)	-2.801*** (0.587)	-3.326*** (0.583)	-3.138*** (0.570)
Adjusted p-value	0.000	0.002	0.001	0.001	0.001
Obs.	1,186	1,186	1,158	1,107	1,107
R ²	0.089	0.082	0.102	0.260	0.307
% of homes in the district whose travel time to district capital is less than 24 hours	-0.0413*** (0.00701)	-0.134*** (0.0269)	-0.102*** (0.0213)	-0.104*** (0.0236)	-0.101*** (0.0229)
Adjusted p-value	0.000	0.000	0.001	0.001	0.001
Obs.	1,186	1,186	1,158	1,107	1,107
R ²	0.033				
Geographic variables	No	No	Yes	Yes	Yes
Population variables	No	No	No	Yes	Yes
Mining	No	No	No	No	Yes

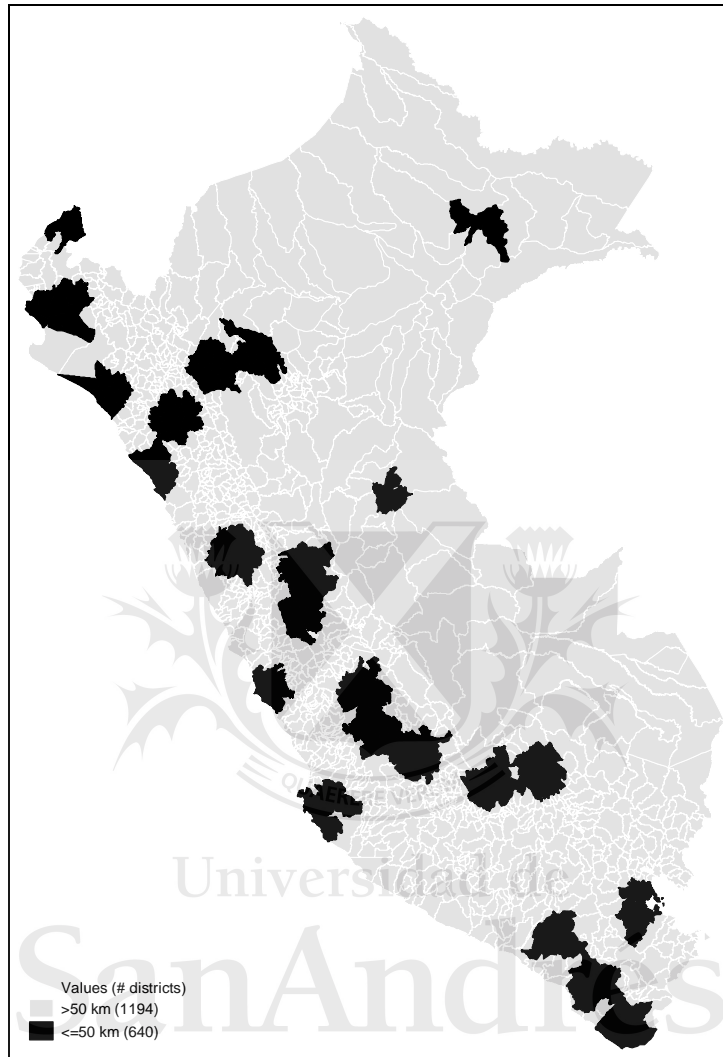
Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Adjusted p-values using Anderson (2007) Multiple Testing Procedure (See in Appendix, Table 16, original p-values). Standard errors clustered at the department level. The parameters of each variable are, respectively, the estimated coefficient/standard error of the estimated parameter and the coefficient of determination of regression R². Note: All the regressions include as regressors the age of the district in years (since its establishment) and the fixed effects for the natural regions of Pulgar Vidal. The geographic variables are the average gradient of the land and the average altitude. In addition, a logarithm of the agricultural area in 1994 has been included. The population variables include the percentage of rural population in 1993, the percentage of the population whose native language was Quechua, Aymara or an Amazonian lowland language in 1993, the percentage of the population with a high school or higher education in 1993, the percentage of the population working in the agricultural sector in 1993, and the logarithm of the population in 1993. Mining is a dummy variable which indicates if the district has mining activity.

SanAndrés

MAP 1
TREE OF MINIMUM DISTANCES BETWEEN DEPARTMENTS
AND THE ROAD NETWORK



MAP 2
SAMPLE OF ESTIMABLE DISTRICTS



APPENDIX

Correlation of the instrument

We analyze the characteristics of the instrument and its correlation with various factors that could explain the districts' agricultural results. One assumption of instrumental variables is that the instrument should affect agricultural output variables only through the road density measurement rather than directly or through other variables. However, this assumption could be violated. Indeed, the population is not randomly distributed in space but rather is strategically distributed according to expectations and possibilities, taking advantage of opportunities offered by the space (van de Valle 2009).

Table 2
CORRELATION OF THE INSTRUMENT WITH SOME REGRESSORS

Control Variables	(1) National
% Rural Population in 1993	-0.0098 (0.1147)
% Population with an indigenous native language in 1993	-0.0004 (0.0010)
% Population with a HS education or more in 1993	-0.0000 (0.0002)
% Population in the agricultural sector in 1993	0.0048 (0.0047)
Ln[Population in 1993]	-0.1645*** (0.0335)
Gradient of the land	-0.0289*** (0.0103)
Altitude (MASL)	0.0000 (0.0001)
1=District with mining	0.2160*** (0.0835)
Ln[Cultivated land in 1993]	-0.0258 (0.0285)
Year district was created	0.0015*** (0.0006)
Constant	12.2708*** (0.3624)
Obs	1,102
R ²	0.117

Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1
Note: All regressions include natural region dummies.

To illustrate this, the previous table presents regressions that evaluate the correlation of the instrument with a set of regressors from 1993 and/or time invariant

variables. As can be seen, households further away from the lines of the Euclidean tree are marked by more mining activity, a lower percentage of people with secondary or higher education, more people working in the agricultural sector, a lower population, and less steeply inclined terrain.

A particularly important variable takes into account how long a district has existed.⁴¹ This variable is significant in the regression and is positively correlated with the instrument. This means that the cities that are more distant to the Euclidean straight line connections between departmental capitals are also the oldest. Thus, given that these cities should also be more developed (or have better agricultural results), the instrument can be correlated directly with agricultural performance. To control for this potential problem, we controlled for the age of districts.

In addition, it is reasonable to expect that the results will vary depending on the natural region where the districts are located. Therefore, fixed effects at the level of the eight natural regions identified by Javier Pulgar Vidal (1938) are included in all regressions.

Robustness

According to the identification strategy, if our instrument is valid then it should be linked to the dependent variables only through road density and not directly, and the relationship should be strong and manifest even in reduced forms; that is, in regressions that directly link outcome variables with the instrumental variable. To verify this, in Table 2, we estimate reduced forms of some variables.

For interpretation of parameters, it is necessary to remember that the instrument is negatively related to road density: the more distant the districts are from the lines of the Euclidian tree, the lower road density. Thus, if the reduced form exhibits a positive relationship between road density and the dependent variable, then the relationship between the instrument and the dependent variable should be negative. According to Table 7, this argument is valid in the case of variables for which significant impacts are estimated; the proportion of sales is negatively correlated with the instrument, but the consumption and own-input are positively correlated. This occurs both in the national sample and the Andean sample. In addition, for the Andean sample, the instrument is negatively correlated with the rate of diversification and per capita income. These

⁴¹ In Peru, new administrative districts may be created within provinces.

results validate the estimates in the previous tables and provide more evidence about the causality.



Universidad de
San Andrés

TABLE2
PRODUCTION DESTINATION: REDUCED FORMS

Variables dependientes	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Use for production				Gross Value of Agricultural Production	Diversification index	HDI	Ln[Per cápita income]	Sufficiency of agricultural income	Diversification of employment
	Sales	Self-consumption	Own-input	Feed for animals						
Ln[Euclidian Distance]	-0.020** (0.008)	0.022*** (0.007)	0.004** (0.002)	-0.006 (0.008)	0.029 (0.030)	-0.021 (0.015)	0.002 (0.003)	-0.013 (0.014)	-0.001 (0.004)	-0.000 (0.006)
Adjusted p-value	0.010	0.002	0.549	0.215						
R ²	0.532	0.392	0.040	0.267	0.646	0.231	0.521	0.443	0.106	0.093
Obs.	1,090	1,090	1,090	1,090	1,088	1,090	1,102	1,102	1,098	1,098
Geographic variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Population variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mining	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Adjusted p-values using Anderson (2007) Multiple Testing Procedure (See in Appendix, Table 14, original p-values). The parameters of each variable are, respectively, the estimated coefficient/standard error of the estimated parameter and the coefficient of determination of regression R². Standard errors clustered at the department level. Note: All the regressions include as regressors the age of the district in years (since its establishment) and the fixed effects for the natural regions of Pulgar Vidal. The geographic variables are the average gradient of the land and the average altitude. In addition, a logarithm of the agricultural area in 1994 has been included. The population variables include the percentage of rural population in 1993, the percentage of the population whose native language was Quechua, Aymara or an Amazonian lowland language in 1993, the percentage of the population with a high school or higher education in 1993, the percentage of the population working in the agricultural sector in 1993, and the logarithm of the population in 1993. Mining is a dummy variable which indicates if the district has mining activity.

San Andrés

TABLE 11
PRODUCTION USES

Variables	(1) OLS	(2) IV	(3) IV	(4) IV	(5) IV
a) National					
Sales	0.163** (0.031)	0.362** (0.135)	0.322** (0.139)	0.379** (0.157)	0.381** (0.160)
R ²	0.501	0.482	0.493	0.512	0.511
P-value	0.000	0.007	0.021	0.021	0.016
Self - Consumption	0.019 (0.031)	-0.271** (0.138)	-0.334** (0.144)	-0.414** (0.160)	-0.424** (0.164)
R ²	0.315	0.264	0.248	0.289	0.285
P-value	0.560	0.048	0.020	0.009	0.009
Own-Input	-0.010 (0.005)	-0.056** (0.027)	-0.066** (0.029)	-0.078** (0.035)	-0.077** (0.036)
R ²	0.030	-	-	-	-
P-value	0.066	0.033	0.020	0.026	0.031
Feed for animals	-0.172*** (0.022)	-0.034 (0.131)	0.078 (0.137)	0.112 (0.146)	0.120 (0.149)
R ²	0.213	0.195	0.194	0.228	0.225
P-value	-	0.798	0.569	0.441	0.419
Obs	1,155	1,155	1,140	1,090	1,090
Geographic variables	No	No	Yes	Yes	Yes
Population variables	No	No	No	Yes	Yes
Mining	No	No	No	No	Yes

Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1. The parameters of each variable are, respectively, the estimated coefficient/standard error of the estimated parameter and the coefficient of determination of regression R². Standard errors clustered at the department level. Note: All the regressions include as regressors the age of the district in years (since its establishment) and the fixed effects for the natural regions of Pulgar Vidal. The geographic variables are the average gradient of the land and the average altitude. In addition, a logarithm of the agricultural area in 1994 has been included. The population variables include the percentage of rural population in 1993, the percentage of the population whose native language was Quechua, Aymara or an Amazonian lowland language in 1993, the percentage of the population with a high school or higher education in 1993, the percentage of the population working in the agricultural sector in 1993, and the logarithm of the population in 1993. Mining is a dummy variable which indicates if the district has mining activity.

TABLE 12
PRODUCTION (IMPACT ON THE GROSS VALUE OF AGRICULTURAL
PRODUCTION AND DIVERSIFICATION)

Variables	(1) OLS	(2) IV	(3) IV	(4) IV	(5) IV
Diversification index	0.114** (0.062)	0.891** (0.294)	0.435 (0.275)	0.400 (0.281)	0.392 (0.288)
R ²	0.068	-	0.149	0.199	0.200
P-value	0.049	0.002	0.114	0.155	0.173
Obs.	1,155	1,155	1,140	1,090	1,090
Ln[Gross Value of Agricultural production 2009]	0.370** (0.152)	1.998** (0.702)	-0.141 (0.490)	-0.429 (0.556)	-0.548 (0.564)
R ²	0.251	0.183	0.633	0.635	0.635
P-value	0.012	0.004	0.773	0.440	0.331
Obs.	1,151	1,151	1,139	1,088	1,088
Geographic variables	No	No	Yes	Yes	Yes
Population variables	No	No	No	Yes	Yes
Education and Mining	No	No	No	No	Yes

Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1. The parameters of each variable are, respectively, the estimated coefficient/standard error of the estimated parameter and the coefficient of determination of regression R². Standard errors clustered at the department level. Note: All the regressions include as regressors the age of the district in years (since its establishment) and the fixed effects for the natural regions of Pulgar Vidal. The geographic variables are the average gradient of the land and the average altitude. In addition, a logarithm of the agricultural area in 1994 has been included. The population variables include the percentage of rural population in 1993, the percentage of the population whose native language was Quechua, Aymara or an Amazonian lowland language in 1993, the percentage of the population with a high school or higher education in 1993, the percentage of the population working in the agricultural sector in 1993, and the logarithm of the population in 1993. Mining is a dummy variable which indicates if the district has mining activity.

TABLE 13
IMPACT ON WELFARE VARIABLES

Dependen variables	(1) OLS	(2) IV	(3) IV	(4) IV	(5) IV
Human Development Index	0.006 (0.010)	-0.079 (0.050)	-0.019 (0.050)	-0.062 (0.049)	-0.042 (0.049)
P-value	0.610	0.113	0.700	0.211	0.390
Obs.	1,175	1,175	1,153	1,102	1,102
Ln[Per capita income]	0.019 (0.055)	-0.023 (0.253)	0.232 (0.263)	0.125 (0.269)	0.249 (0.269)
P-value	0.701	0.929	0.378	0.642	0.354
Obs.	1,175	1,175	1,153	1,102	1,102
Sufficiency of agricultural income	0.025 (0.019)	0.083 (0.067)	0.012 (0.068)	0.025 (0.079)	0.024 (0.080)
P-value	0.173	0.218	0.861	0.752	0.764
Obs.	1,167	1,167	1,149	1,098	1,098
Diversification of employment	-0.010 (0.024)	-0.095 (0.094)	-0.004 (0.098)	-0.000 (0.107)	0.008 (0.109)
P-value	0.863	0.313	0.966	0.997	0.938
Obs.	1,167	1,167	1,149	1,098	1,098
Geographic variables	No	No	Yes	Yes	Yes
Population variables	No	No	No	Yes	Yes
Education and Mining	No	No	No	No	Yes

Robust standard errors are in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. For each variable, the parameters are, respectively, the estimated coefficient/standard error of the estimated parameter and the number of observations in each regression. Standard errors clustered at the department level. Note: All the regressions include as regressors the age of the district in years (since its establishment) and the fixed effects for the natural regions of Pulgar Vidal. The geographic variables are the average gradient of the land and the average altitude. In addition, a logarithm of the agricultural area in 1994 has been included. The population variables include the percentage of rural population in 1993, the percentage of the population whose native language was Quechua, Aymara or an Amazonian lowland language in 1993, the percentage of the population with a high school or higher education in 1993, the percentage of the population working in the agricultural sector in 1993, and the logarithm of the population in 1993. Mining is a dummy variable which indicates if the district has mining activity.

TABLE 14
PRODUCTION USES: REDUCED FORMS

Dependent variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Use for Production				Gross Value of Agricultural Production	Diversification index	HDI	Ln[Per cápita income]	Sufficiency of agricultural income	Diversification of employment
	Sales	Self-consumption	Own-input	Feed for animals						
Ln[Euclidian Distance]	-0.020** (0.008)	0.022*** (0.007)	0.004** (0.002)	-0.006 (0.008)	0.029 (0.030)	-0.021 (0.015)	0.002 (0.003)	-0.013 (0.014)	-0.001 (0.004)	-0.000 (0.006)
P-value	0.012	0.002	0.023	0.403						
R ²	0.532	0.392	0.040	0.267	0.646	0.231	0.521	0.443	0.106	0.093
Obs.	1,090	1,090	1,090	1,090	1,088	1,090	1,102	1,102	1,098	1,098

Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1. The parameters of each variable are, respectively, the estimated coefficient/standard error of the estimated parameter and the coefficient of determination of regression R². Standard errors clustered at the department level. Note: All the regressions include as regressors the age of the district in years (since its establishment) and the fixed effects for the natural regions of Pulgar Vidal. The geographic variables are the average gradient of the land and the average altitude. In addition, a logarithm of the agricultural area in 1994 has been included. The population variables include the percentage of rural population in 1993, the percentage of the population whose native language was Quechua, Aymara or an Amazonian lowland language in 1993, the percentage of the population with a high school or higher education in 1993, the percentage of the population working in the agricultural sector in 1993, and the logarithm of the population in 1993. Mining is a dummy variable which indicates if the district has mining activity.

TABLE 15
PRODUCTION USES: HETEROGENOUS IMPACTS

Variables dependientes	Small Producers			Medium-Sized Producers			Large producers		
	(1) OLS	(2) IV	(3) IV	(4) OLS	(5) IV	(6) IV	(7) OLS	(8) IV	(9) IV
Sales	0.149*** (0.031)	0.330*** (0.136)	0.326** (0.162)	0.199*** (0.031)	0.439*** (0.141)	0.440** (0.171)	0.231*** (0.041)	0.556*** (0.161)	0.560*** (0.191)
P-value	0.000	0.015	0.044	0.000	0.002	0.010	0.000	0.001	0.003
Self-consumption	-0.026 (0.033)	-0.312** (0.145)	-0.479*** (0.176)	-0.032 (0.030)	-0.301** (0.140)	-0.459*** (0.167)	-0.013 (0.037)	-0.277** (0.135)	-0.386** (0.159)
P-value	0.424	0.031	0.006	0.279	0.031	0.005	0.736	0.040	0.015
Own-input	-0.006 (0.003)	-0.026 (0.015)	-0.033 (0.019)	-0.005 (0.003)	-0.012 (0.015)	-0.022 (0.017)	-0.002 (0.011)	-0.059 (0.044)	-0.082 (0.055)
P-value	0.064	0.072	0.079	0.126	0.441	0.181	0.849	0.178	0.178
Feed for animals	-0.117*** (0.020)	0.008 (0.110)	0.186 (0.134)	-0.163*** (0.023)	-0.126 (0.122)	0.041 (0.141)	-0.216*** (0.030)	-0.220 (0.156)	-0.092 (0.175)
P-value	0.000	0.945	0.163	-	0.304	0.770	-	0.159	0.601
Obs	1,144	1,144	1,082	1,135	1,135	1,074	1,059	1,059	1,004
Obs	713	713	676	704	704	669	631	631	603
Geographic variables	No	No	Yes	No	No	Yes	No	No	Yes
Population variables	No	No	Yes	No	No	Yes	No	No	Yes
Mining	No	No	Yes	No	No	Yes	No	No	Yes

Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1. For each variable, the parameters are, respectively, the estimated coefficient and the associated standard error. Standard errors clustered at the department level. Note: All the regressions include as regressors the age of the district in years (since its establishment) and the fixed effects for the natural regions of Pulgar Vidal. The geographic variables are the average gradient of the land and the average altitude. In addition, a logarithm of the agricultural area in 1994 has been included. The population variables include the percentage of rural population in 1993, the percentage of the population whose native language was Quechua, Aymara or an Amazonian lowland language in 1993, the percentage of the population with a high school or higher education in 1993, the percentage of the population working in the agricultural sector in 1993, and the logarithm of the population in 1993. Mining is a dummy variable which indicates if the district has mining activity.

TABLE 16
TRAVEL TIME

Dependent Variable	(1)	(2)	(3)	(4)	(5)
	OLS	IV	IV	IV	IV
Travel time from home to district capital	-1.740*** (0.147)	-2.225*** (0.591)	-2.801*** (0.587)	-3.326*** (0.583)	-3.138*** (0.570)
P-value	0.000	0.002	0.000	0.000	0.000
Obs.	1,186	1,186	1,158	1,107	1,107
R ²	0.089	0.082	0.102	0.260	0.307
% of homes in the district whose travel time to district capital is less than 24 hours	-0.0413*** (0.00701)	-0.134*** (0.0269)	-0.102*** (0.0213)	-0.104*** (0.0236)	-0.101*** (0.0229)
Adjusted p-value	0.000	0.000	0.001	0.001	0.001
Obs.	1,186	1,186	1,158	1,107	1,107
R ²	0.033				
Geographic variables	No	No	Yes	Yes	Yes
Population variables	No	No	No	Yes	Yes
Mining	No	No	No	No	Yes

Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1. The parameters of each variable are, respectively, the estimated coefficient/standard error of the estimated parameter and the coefficient of determination of regression R². Standard errors clustered at the department level. Note: All the regressions include as regressors the age of the district in years (since its establishment) and the fixed effects for the natural regions of Pulgar Vidal. The geographic variables are the average gradient of the land and the average altitude. In addition, a logarithm of the agricultural area in 1994 has been included. The population variables include the percentage of rural population in 1993, the percentage of the population whose native language was Quechua, Aymara or an Amazonian lowland language in 1993, the percentage of the population with a high school or higher education in 1993, the percentage of the population working in the agricultural sector in 1993, and the logarithm of the population in 1993. Mining is a dummy variable which indicates if the district has mining activity.

San Andrés