Evidence from a Natural Experiment on the Development Impact of Windfall Gains:

The Camisea Fund in Peru

Leonardo Corral
Heath Henderson
Juan Jose Miranda

Evidence from a Natural Experiment on the Development Impact of Windfall Gains:

The Camisea Fund in Peru

Leonardo Corral
Heath Henderson
Juan Jose Miranda

Inter-American Development Bank
Drake University
The World Bank

http://www.iadb.org

OPTIONAL: Type address for correspondence

OPTIONAL: Type Authors name and eMail

Evidence from a Natural Experiment on the Development Impact of Windfall Gains: The Camisea Fund in Peru*

Leonardo Corral
Office of Strategic Planning and Development Effectiveness
Inter-American Development Bank

Heath Henderson[†]
College of Business and Public Administration
Drake University

Juan Jose Miranda
Environment and Natural Resources Global Practice
The World Bank

February 24, 2016

Abstract

We study the economic effect of windfall gains by examining a Peruvian natural experiment. The Camisea Fund for Socioeconomic Development (FOCAM) is an inter-governmental fiscal transfer scheme that allocates natural gas royalties generated by the Camisea Gas Project to eligible subnational governments. We exploit the rules governing FOCAM allocation to identify the effect of the transfers on municipal accounts, local infrastructure, and economic development. Using a newly constructed district-level dataset for the years 2005 and 2012, we find evidence of positive impacts on municipal capital expenditures and local infrastructure. However, we also find evidence of a negative impact on municipal current expenditures. More specifically, we find that municipalities with low absorptive capacity coped with the increased administrative burden of FOCAM transfers by reallocating administrative effort toward (away from) executing capital (current) expenditures.

Key words: Inter-governmental Fiscal Transfers; Natural Experiment; Peru; Windfall Gains
JEL codes: C14; H70; O10; O54; Q32

1 Introduction

The economic effect of windfall gains (e.g. foreign aid, natural resource rents, and inter-governmental fiscal transfers) remains controversial, due in part to identification issues in the empirical literature.[1] Consider, for example, the empirical literature on the effect of foreign aid on economic growth. Using 2SLS in the context of cross-country growth regressions, Burnside and Dollar (2000) found that aid only positively affected growth in developing countries with sound fiscal, monetary, and trade policies. Hansen and Tarp (2001), however, found an unconditionally positive aid-growth relationship when additionally accounting for endogeneity bias due to country-specific effects. Rajan and Subramanian (2008) questioned the instrumentation strategies used in prior cross-country studies and alternatively modeled the supply of aid based on donor (as opposed to recipient) characteristics. Contrary to the aforementioned studies, Rajan and Subramanian found little evidence of an aid-growth relationship, but their instrumentation strategy itself was questioned by Bazzi and Clemens (2013).[2]

The empirical literature on the economic effect of natural resource rents has also encountered identification issues. Also using cross-country growth regressions, Sachs and Warner (1995) found an inverse relationship between economic growth and the (initial) ratio of natural resource exports to GDP. Brunnschweiler and Bulte (2008), however, argued that Sachs and Warner's resource dependence measure was endogenous to underlying structural factors and became statistically insignificant when instrumented. The authors then contended that a measure of resource abundance (i.e. subsoil assets) was potentially exogenous, and found a positive relationship between resource abundance and economic growth in a cross-country setting. Van der Ploeg and Poelhekke (2010) nevertheless suggested that Brunnschweiler and Bulte's measure of subsoil assets was itself endogenous given its proportionality to resource rents. The authors then found that the subsoil assets variable became statistically insignificant when instrumented, though they cautioned that their instrument was also "somewhat endogenous" (pg. 47).

In light of such identification issues, van der Ploeg (2011) suggested that "[t]he road forward might be to exploit variation within a country where . . . the danger of spurious correlation is minimized" (pg. 381). While research at the subnational (or regional) level may offer additional insights into the economic effect of foreign aid or natural resource rents, it is also relevant to understanding another increasingly important

---

[1]We follow Dalgaard and Olsson (2008) and define windfall gains by their disproportionate revenue-to-cost ratio as compared to revenues from the standard production of goods and services. See Arndt et al. (2010) for a review of the foreign aid literature and van der Ploeg (2011) for a review of the literature on natural resource rents.

[2]Bazzi and Clemens argued that Rajan and Subramanian were effectively instrumenting foreign aid with recipient population size, even though Rajan and Subramanian themselves stated that such an instrument was unlikely to satisfy the exclusion restriction.

type of windfall gain: inter-governmental fiscal transfers. Recent decades have seen a movement toward fiscal decentralization in many transition and developing economies (De Mello 2000; Arzaghi and Henderson 2005).[3] As the objective of decentralization is typically to improve allocative efficiency, distributional equity, or macroeconomic growth/stability, complex systems of inter-governmental transfers commonly accompany decentralization efforts (Bird 1993; Bird and Smart 2002). Subnational governments in developing countries are highly and increasingly dependent on these transfers (Gadenne and Singhal 2013),[4] but relatively little is known about their economic effect (Paler 2011; Becker et al. 2013).

Paler (2011) discussed the theoretical mechanisms through which windfall gains can affect economic outcomes at the subnational level. First, Dutch disease can operate at the subnational level if the inflow of capital induces an increase in the local price level. While the windfall can provide local governments with revenue to fund long-run development, rising factor prices can also squeeze segments of the tradable-goods sector and hamper productivity growth. Second, windfall gains can strengthen or undermine local governing institutions. Strengthening can occur when additional resources facilitate capacity building and undermining can occur when windfalls induce corruption or rent-seeking behavior. Finally, windfall gains can hamper economic growth by inciting local conflict, particularly by exacerbating grievances due to social inequality, forced migration and environmental degradation, or ethnic tensions.

While it is evident that the economic effect of subnational windfall gains is theoretically indeterminate, the empirical evidence also remains inconclusive. Caselli and Michaels (2009), for example, exploited variation in oil output across Brazilian municipalities to examine the economic effect of natural resource windfalls. The authors found little evidence of Dutch disease-type effects or improvements in local living standards, even though oil abundance caused municipal revenues and spending to increase. Further examining Brazilian municipalities, Brollo et al. (2010) used regression discontinuity design to analyze the effect of federal transfers on political corruption. Consistent with anecdotal evidence discussed in Caselli and Michaels, the authors found that larger transfers increased political corruption and reduced the quality of candidates for mayor.[5] While the Brazilian evidence suggests that windfall gains can undermine local governing institutions, positive effects have been witnessed in other contexts.

Becker et al. (2013), for example, analyzed the economic effect of a regional transfer scheme in the European Union. Also using regression discontinuity design, the authors found that those regions with

---

[3]Dillinger (1994) claimed that 63 out of the 75 developing and transition economies with populations greater than 5 million people have embarked or are embarking on some form of fiscal or political decentralization.

[4]More specifically, Gadenne and Singhal (2013) document that the average share of non-tax revenues in subnational revenues (i.e. the average fiscal gap) increased in developing countries from 49 percent to 62 percent between 1996-2000 and 2006-2010.

[5]See Vicente (2010) for further evidence suggesting that natural resource rents tend to induce corruption.

sufficient human capital and relatively high-quality institutions witnessed faster per capita income growth. Exploiting an exogenous upsurge in coca prices and cultivation in Colombia, Angrist and Kugler (2008) found some positive economic effects (e.g. increased self-employment earnings) of windfall gains as well. While the authors further witnessed increased violence in regions where coca cultivation increased, Dube and Vargas (2013) found that price increases for other Colombian exports (e.g. coffee) can mitigate conflict by reducing labor supplied to violent resource appropriation. Given the evident lack of resolution in the existing research, we contend that further empirical work is needed to better understand the effect of windfall gains.

We thus examine the economic effect of windfall gains by exploiting a Peruvian natural experiment. Operational in 2004, the Camisea Gas Project – a natural gas extraction and distribution project – is one of Peru's largest energy infrastructure projects. Largely through its two natural gas pipelines, the project impacts six of Peru's 25 regions as well as the province of Lima. To support the economic, social, and environmental development of the regions affected by the pipelines, the Peruvian government established the Camisea Fund for Socioeconomic Development (FOCAM). FOCAM is an inter-governmental fiscal transfer scheme that allocates a percentage of the natural gas royalties received by the central government to eligible subnational governments. We exploit the rules governing FOCAM allocation to identify the effect of the transfers on municipal accounts, local infrastructure, and economic development.

Using a newly constructed district-level dataset for the years 2005 and 2012, we first find that treated municipalities witnessed statistically significant increases in capital expenditures. Given that FOCAM transfers are earmarked for capital expenditures, this result is expected. Second, we find that treated districts also witnessed statistically significant positive impacts on local infrastructure. For example, we find that treated districts had greater access to internet and constructed more primary roads than their control group counterparts. Finally, we unexpectedly find evidence of a negative impact on municipal current expenditures. Anecdotal and quantitative evidence suggests that capacity constraints (e.g. lack of qualified technical staff) have limited budget execution rates, and we find that low capacity municipalities coped with the increased administrative burden of FOCAM transfers by reallocating administrative effort toward (away from) executing capital (current) expenditures.

In what follows, Section 2 provides background information on the Camisea Gas Project and the associated FOCAM transfers. Section 3 discusses the available data, Section 4 outlines our identification strategy, and Section 5 presents the results of our econometric analysis. Finally, Section 6 provides discussion and concluding remarks.

2   The Camisea Gas Project

In this section, we first provide a brief overview of Peru's recent decentralization efforts so as to contextualize the Camisea Gas Project. We then turn to discussing the Camisea Gas Project in detail, after which we describe the rules governing allocation of the Camisea Fund for Socioeconomic Development (FOCAM). Finally, we conclude the section by arguing that receipt of FOCAM transfers is exogenous to select regions.

Approval and promulgation of the Constitutional Reform Law of 2002 launched the decentralization process in Peru. The reform established that the territory of Peru was to consist of regions, provinces, and districts. Specifically, the country was divided into 26 units: 25 regions and the province of Lima. The regions were subdivided into provinces, and all provinces were subdivided into districts. In each jurisdiction, national, regional, and local governments were constituted and organized. The constitutional reform added that the sphere of regional governments was the regions and that the sphere of local governments was the provinces, districts, and towns. Several further legislative acts were also approved to promote administrative and fiscal decentralization, and to enhance citizen participation, particularly in yearly budgeting exercises. In addition, in order to optimize the use of public resources for investment, the National System for Public Investment (SNIP) was extended to regulate all regional and local governments (Contraloría General de la República del Perú 2014).

Loayza et al. (2014) documented that these rapid changes overwhelmed many municipalities. On the one hand, decentralization transferred additional responsibilities and resources from the central government to regional and local governments. On the other hand, strong fiduciary requirements and demanding budgeting guidelines made budgeting and execution more difficult. Many municipalities have thus struggled to spend their allocated budget. In 2009, for example, municipalities spent on average 74 percent of their budgeted expenditures. Further, current expenditures tended to be executed at better rates than capital expenditures (83 percent versus 71 percent, respectively). Anecdotal and quantitative evidence suggests that capacity constraints (e.g. lack of qualified technical staff) have limited execution rates considerably, particularly for capital expenditures (see Loayza et al. [2014] for further information). It is in this context that the Camisea Gas Project was implemented.

The Camisea Gas Project is Peru's biggest energy project. Located in the department of Cusco, the Camisea natural gas fields have proven reserves of 9 trillion cubic feet, among the largest in Latin America. Royal Dutch/Shell started exploring the fields in the mid-1980s, but walked away in 1998 after contractual disputes with the Government of Peru (GoP) (The Economist 2003). Peru's Private Investment Commission (COPRI) then issued an international call for proposals to further develop the fields. The Camisea Gas

Project had three primary components: (1) exploration of the Camisea gas fields (Block 88); (2) transportation of the natural gas to Peru's coast via pipeline; and (3) distribution of the gas in Lima and Callao. The project also involved construction of a gas liquification plant and a terminal for exporting natural gas liquids in the southern part of the country (Inter-American Development Bank 2002).

In 2000, the GoP awarded a license to develop the fields to an "upstream" consortium headed by Pluspetrol of Argentina. A second "downstream" project, led by the consortium *Transportadora de Gas del Peru* (TGP), consisted of an agreement to build, own, and operate two major pipeline systems: a 700 kilometer natural gas pipeline and a 540 kilometer natural gas liquids (NGL) pipeline (see Figure 1). The two pipelines were to be laid in parallel trenches on a common right-of-way extending from a gas processing plant at Las Malvinas to an NGL processing and shipping facility near the port of Pisco.[6] The natural gas pipeline was to continue to Lima from a point east of Pisco. Across all components (i.e. including extraction, transportation, and distribution of the natural gas), the total estimated cost of the Camisea Gas Project was $1.45 billion (Inter-American Development Bank 2002; The Economist 2003).

To finance the pipeline, TGP requested a $75 million loan from the Inter-American Development Bank (IDB). Due to the Camisea field location, deep in one of the more pristine parts of the Peruvian jungle, the IDB was concerned about creating benefits for local populations and safeguarding the environment. As such, strict environmental and social protocols were designed and enforced by the IDB in the layout and construction of the pipeline. The IDB also provided two loans for the project: the $75 million loan to TGP to help finance the downstream component and a $5 million loan to the GoP. The loan to the GoP was intended to serve two purposes: (1) to help strengthen governmental capacity to supervise, monitor, and inspect the project's environmental and social aspects; and (2) to help carry out actions to create sustainable and balanced development in the project's area of influence (Inter-American Development Bank 2002).

As part of the loan negotiations, the GoP presented the IDB with a letter of "XXI Commitments" outlining specific actions. Commitment III called for the GoP to establish a fund – financed by natural gas royalties – to support the economic, social, and environmental development of the areas affected by the pipeline. As a result, the Camisea Fund for Socioeconomic Development (FOCAM) was created by law in December of 2004. All regional and local governments in the regions of Ayacucho, Huancavelica, Ica, and Lima (excluding metropolitan Lima) were designated as recipients of FOCAM transfers (see Figure 1) (Inter-American Development Bank 2004; Ministerio de Economía y Finanzas 2015). The region of Ucayali was subsequently added following a series of demonstrations over the use of local waterways by the project.

---

[6]The Las Malvinas plant is located in the Ucayali Basin in the region of Cusco just over 400 kilometers east of Lima. Pisco is approximately 200 kilometers south of Lima.

Importantly, the region of Cusco – where the Camisea fields are located – was not designated as a beneficiary of FOCAM transfers as they were to benefit from a separate transfer scheme called the Gas Canon (Munilla 2010).

By law, 25 percent of the GoP's natural gas royalties are designated for FOCAM, though only after Gas Canon and other deductions are made. For Ayacucho, Huancavelica, Ica, and Lima, FOCAM funds are distributed as follows:

- Regional governments receive 30 percent of FOCAM funds, distribution across which depends on population, unsatisfied basic needs (UBN), and the length of the pipeline in each jurisdiction;

- Provincial governments receive 30 percent of FOCAM funds, distribution across which depends on population and UBN in each jurisdiction;

- District governments crossed by the pipeline receive 15 percent of FOCAM funds, distribution across which depends on population, UBN, and the length of the pipeline in each jurisdiction;

- District governments not crossed by the pipeline receive 15 percent of FOCAM funds, distribution across which depends on population and UBN in each jurisdiction;

- Public universities in the regions of Ayacucho, Huancavelica, Ica, and Lima receive an equal share of 10 percent of FOCAM funds.

Given the unique circumstance, Ucayali was designated separate funds amounting to 2.5 percent of the GoP's share of royalties. For all beneficiaries, the GoP places spending requirements on FOCAM funds and the resources are earmarked for capital rather than current expenditures.[7] The Camisea project began operating in 2004 and the first disbursement from FOCAM was in 2005 (Munilla 2010; Ministerio de Economía y Finanzas 2015).

FOCAM transfers represent a natural experiment, where we define natural experiment as "a transparent exogenous source of variation in the explanatory variables that determine the treatment assignment" (Meyer 1995, pg. 151). In construction of the pipeline, TGP naturally sought to minimize pipeline length by taking the most direct route from the processing plant at Las Malvinas to the shipping facility near Pisco, and then to Lima from east of Pisco. Nevertheless, other considerations were also important. For example, TGP attempted to avoid rugged topography, minimize water crossings, and bypass fragile habitats and protected areas (Walsh Peru S.A. 2002). Municipalities were unable to influence the path of the pipeline and were thus limited in their ability to influence their receipt of FOCAM funds. Accordingly, conditional on key control variables (see Sections 3 and 4), we contend that receipt of FOCAM transfers is exogenous, albeit only to select regions.

---

[7]Specifically, the funds must be devoted to "social and economic infrastructure" (Ministerio de Economía y Finanzas 2015).

So as to proceed conservatively, we drop the regions of Ucayali and Lima (including the province of Lima) from our econometric analysis. While Ucayali is not located along the pipeline, the region was added to the list of FOCAM-receiving regions following a series of demonstrations over the use of local waterways by the Camisea Gas Project. Given this self-selection issue, we believe it is preferable to drop Ucayali from the analysis. Lima is dropped from the analysis for two reasons: given the size and nature of the local economy, we do not expect (1) FOCAM transfers to have a substantial effect and (2) to be able to identify reasonable counterfactuals for the associated municipalities. Finally, note that Cusco, the producing region, is also not included in the analysis. While Cusco does not receive FOCAM transfers, the region does receive Gas Canon transfers. As there are likely some endogenous factors that influenced the exploration and development of the Camisea fields, we believe it is also preferable to drop Cusco from the analysis. Thus, our "treated" units are municipalities in the regions of Ayacucho, Huancavelica, and Ica (see Figure 1).

3 Data

In this section, we discuss the available data, which permits analysis at both the region and district levels. After discussing our data sources, we use the region level data to (1) better understand the magnitude of the FOCAM transfers and (2) assess our "parallel trends" assumption.[8] Region level data is used for these exercises because each benefits from time series data for select variables (e.g. GDP), which is only available at the region level. The unit of our econometric analysis, however, is the district, so we conclude the section by discussing the district-level data and examining descriptive statistics.

We draw upon five data sources in this and subsequent sections. First, the *Sistema de Información Regional para la Toma de Decisiones* (SIRTOD) is a region-level database that provides economic indicators, education and health statistics, and demographic information on an annual basis. Second, the *Registro Nacional de Municipalidades* (RENAMU) is an annual census of local governments that collects information on infrastructure, public services, and economic activity in each jurisdiction. Third, the *Sistema de Integración Contable de la Nación* (SICON) is a database containing annual information on public sector accounts and provides information on transfers received by regional and local governments. Fourth, the *Censo de Población y Vivienda* (CPV) is a nationwide census that periodically collects demographic and socio-economic information at the individual and household level.

---

[8]The parallel trends assumption states that the treated group would have witnessed a trend similar to the control group if the treated group had not been treated.

Finally, as a proxy for economic activity, we use nighttime lights data from the Defense Meteorological Satellite Program's Operational Linescan System (DMSP-OLS).[9] The data is processed by the National Geophysical Data Center (NGDC) and is available on the National Oceanic and Atmospheric Administration's (NOAA) website. The intensity of light in each pixel is reported as an integer between 0 (no light) and 63. In order to facilitate interannual comparisons, the data are calibrated to the year 1999 to eliminate satellite instrumental variations. The total amount of nighttime light in each administrative unit (i.e. regions or districts) is calculated as the sum across all pixels within that unit.[10] The geographical boundaries for the administrative units are downloaded from the World Agroforestry Centre's (ICRAF) Landscapes Portal. For a full discussion of the nighttime lights data, including discussion of the well-known issues of saturation and overglow, see Chen and Nordhaus (2011), Henderson et al. (2012), or Michalopoulos and Papaioannou (2013).

To get an understanding of the magnitude of the FOCAM transfers, we use SIRTOD and SICON data to calculate (1) FOCAM transfers as a percentage of government revenue and (2) FOCAM transfers as a percentage of GDP. We calculate these quantities for each of the five recipient regions for the years 2005-2012 and the results are presented in Figure 2.[11] With the exception of Lima, we see that the FOCAM transfers are non-negligible, hovering around 10-15 percent of government revenue and 1-1.5 percent of GDP. Further, following trends in the production value of natural gas, the magnitude of the transfers tends to increase as time passes, reaching a peak in 2011. In Huancavelica, for example, FOCAM transfers reach a height of approximately 2.25 percent of GDP and 20 percent of government revenue in 2011. The magnitudes of the transfers thus suggest the potential for impact on local economic development.

To identify the impact of the FOCAM transfers we rely on the well-known parallel trends assumption. To assess the plausibility of this assumption, we graphically compare the *ex ante* trends of select variables between the treated and donor pool regions.[12] While time series data of a sufficient length is scarce, we are able to undertake this exercise at the region level using two series: GDP per capita (from the SIRTOD database) and nighttime lights per capita (from the DMSP-OLS database). Using population-weighted

---

[9]Previous studies have validated the use of nighttime lights data to predict economic activity. For example, see Henderson et al. (2012).

[10]When a pixel is shared by different administrative units, the proportion of light within a given unit is extracted and assigned to that unit.

[11]The GDP and government revenue data are from the SIRTOD database and the FOCAM data is from the SICON database. The FOCAM transfer (government revenue) variable is calculated as the sum of FOCAM transfers (government revenue) for all sub-national governments within a given region. Regarding the chosen years, recall that 2005 is the first year that FOCAM funds were transferred to sub-national governments. Further, 2012 is the most recent year for which we have data. The unit of analysis here is the region because some data (e.g. GDP) is only available at the region level.

[12]We also tested the parallel trends assumption statistically. Using the regional data, this entailed running difference-in-differences regressions for only the pre-treatment period. We undertook this exercise for a number of alternative dependent variables and in no case did we find evidence that the parallel trends assumption is violated. Results are available upon request.

8

averaging to aggregate across the constituent regions in each group, Figure 3 plots GDP per capita and nighttime lights per capita for the treated and donor pool groups.[13] Note that the treated group includes Ayacucho, Huancavelica, and Ica whereas the donor pool group includes all regions that did not receive FOCAM transfers (excluding Cusco).[14] It is evident from Figure 3 that the pre-2005 (or pre-FOCAM) trends for the treated and donor pool regions are nearly perfectly parallel in both panels.[15]

We now turn to the district-level data that we use in our econometric analysis. The district is taken as the unit of analysis because it is the smallest administrative division that receives FOCAM transfers. We have data at the district level for the years 2005 and 2012, which we use as the *ex ante* and *ex post* periods in our econometric specification (see Section 4). Recall that FOCAM transfers began in 2005 so it is natural to use this year as our *ex ante* period. The year 2012 is used as the *ex post* period because it is the most recent year for which we have data.[16] While Peru currently has 1,838 districts, we only observe 1,824 districts across the two years because some of the current districts did not exist in 2005. Further, we drop the districts in regions for which treatment is endogenous (see Section 2) and any district with unreliable and/or missing data. Our final sample size is thus 1,492 (246 are treated districts and 1,246 are donor pool districts).

Table 1 presents definitions for all variables used in the analysis. With the exception of the nighttime lights variable, which is from the DMSP-OLS database, all dependent variables are from the RENAMU database. The choice of dependent variables follows from our interest in the effect of FOCAM transfers on municipal accounts (i.e. *expenditure*, *capital*, and *current*), local infrastructure (e.g. *internet*, *roads*, and *water*), and economic development (i.e. *lights*). The independent variables come from a variety of sources, including the RENAMU, SICON, CPV, and DMSP-OLS databases. We include variables to control for initial development (i.e. *lights*), geographical location (i.e. *highlands*, *coast*, and *area*), sociodemographic attributes (i.e. *population*, *dependents*, and *literacy*), and local government characteristics (i.e. *transfers* and *staff*). As will be seen, *staff* is a particularly important variable as we use it to proxy municipal absorptive capacity (Loayza et al. 2014).

---

[13]Nighttime lights per capita is defined here as the total amount of nighttime light in each region divided by the region's population.

[14]As discussed in Section 2, treatment is endogenous to Cusco, Lima, and Ucayali so these regions are dropped from the treated group.

[15]Two other aspects of Figure 3 are worth noting. First, for GDP per capita in particular, there exists non-negligible *ex ante* differences in variable levels across treated and donor pool regions. Second, around 2008, GDP per capita in the treated regions began to grow more rapidly than in the donor pool regions whereas, paradoxically, the opposite is true for nighttime lights per capita. These two considerations raise important questions about identifying the impact of FOCAM transfers through a simple difference in means and thereby further motivate our econometric analysis.

[16]We made an attempt to construct a panel dataset that used interim years, but could not do so because of data compatibility issues. Specifically, a major constraint was that RENAMU questionnaires varied on a yearly basis, which made it difficult to construct comparable variables across interim years.

9

To better understand the *ex ante* differences between treated and donor pool districts, Table 2 presents select descriptive statistics for the independent variables for the year 2005. The second and third columns present variable means for the donor pool and treated districts, respectively. The final three columns present normalized difference statistics for the pre-matching (ND-PM) and post-matching (ND-SS and ND-KM) samples, respectively.[17] (We reserve discussion of the post-matching samples until Section 5.) Looking at the pre-matching normalized differences, the most substantive differences occur in the sociodemographic variables (i.e. *population*, *dependents*, and *literacy*), which all have normalized differences in excess of 0.25 in absolute value. Notable differences are, however, not confined to these variables as *highlands*, *staff*, and *lights* also witness non-negligible normalized differences. As linear regression methods tend to be sensitive to the specification when normalized differences exceed one quarter (Imbens and Wooldridge 2009), the next section pays particular attention to this extrapolation issue in discussing our identification strategy.

4 Methodology

In this section, we consider our strategy to identify the impact of FOCAM transfers. As treatment status is arguably exogenous to the recipient district (see Section 2), we seek to estimate the "average treatment on the treated" (ATT). In the below discussion of our identification strategy, we rely heavily on the work of King and Zeng (2006) as well as Iacus et al. (2012). The reader is referred to those works for a fuller discussion.

Let T denote the treatment variable where $T = 1$ if a district receives FOCAM transfers and $T = 0$ if a district does not receive FOCAM transfers. Further, let Y denote the dependent variable (further discussed below) and define the causal effect of FOCAM transfers on the dependent variable as a function of potential outcomes. That is, let Y_1 denote the outcome of the dependent variable in the presence of FOCAM transfers and Y_0 denote the outcome in the absence of FOCAM transfers. For a given district, Y_1 and Y_0 cannot then be simultaneously observed. As mentioned, we are primarily interested in the ATT, which is defined as follows:

(1) $$\tau = E(Y_1|T = 1) - E(Y_0|T = 1)$$

---

[17]The normalized difference statistic is calculated as $(\bar{X}_1 - \bar{X}_0)/\sqrt{S_0^2 + S_1^2}$ where $\bar{X}_w$ denotes the sample mean and S_w^2 denotes the sample variance for the treated ($w = 1$) and donor pool ($w = 0$) districts. The normalized difference, as opposed to the t-statistic, provides a way to assess covariate imbalance that is scale and sample size free.

where the first term is observable as Y_1 is observable for the treated districts and the second term is unobservable as Y_0 is not observed for the treated districts. It is thus impossible to directly observe the causal effect and this is the "fundamental problem of causal inference" (Holland 1986).

Consider then an estimator of τ – denoted by $\hat{\tau}$ – as the coefficient on T from a regression of Y on a constant and T. King and Zeng (2006) derived the following decomposition of the bias of $\hat{\tau}$ as an estimator of τ:

$$
(2) \qquad\qquad E(\hat{\tau} - \tau) = \Delta_o + \Delta_p + \Delta_i + \Delta_e
$$

where Δ_o denotes omitted variable bias, Δ_p denotes post-treatment bias, Δ_i denotes interpolation bias, and Δ_e denotes extrapolation bias. Omitted variable bias, Δ_o, occurs when variables that are correlated with both treatment status and the dependent variable are omitted from the regression. Post-treatment bias, Δ_p, is the result of controlling for variables that are themselves a consequence of treatment. Interpolation bias, Δ_i, is due to a failure to properly adjust for independent variables within the observed range of the data. Finally, extrapolation bias, Δ_e, results when there is a failure to properly adjust for independent variables when extrapolating beyond the observed range of the data.

The above bias decomposition identifies the potential issues associated with estimating τ using observational data and permits us to explicitly address the assumptions that accompany our identification strategy. To this end, let X denote a vector of control variables, the appropriate choice of which enables us to reasonably make three simplifying assumptions. First, King and Zeng (2006) demonstrate that if

$$
(3) \qquad\qquad E(Y_0|T = 1, X) = E(Y_0|T = 0, X)
$$

holds, then the first component of Eq. (2) vanishes (i.e. $\Delta_o = 0$).[18] Given the natural experiment (see Section 2) and a rich set of control variables (see Section 3), we contend that assuming $\Delta_o = 0$ is indeed reasonable. Second, if those characteristics included in X include only pre-treatment characteristics (i.e. characteristics that are not a consequence of treatment status), then we may also reasonably assume $\Delta_p = 0$. Finally, as interpolation bias arises when controlling for X using the wrong functional form (e.g. linear rather than quadratic), with the appropriate regression diagnostics we may also assume $\Delta_i = 0$.

---

[18]That is, if the appropriate set of control variables is included, treatment assignment is random and omitted variable bias is eliminated.

With the above assumptions, Eq. (2) becomes $E(\hat{\tau} - \tau) = \Delta_e$ and we now turn to our strategy to mitigate this extrapolation bias. Extrapolation bias occurs when some members of the treated (control) group witness certain values of X with a positive probability that no members of the control (treated) group witness. Matching methods are data pre-processing algorithms that can serve to mitigate such extrapolation bias by discarding observations outside the region of common support. To this end, we use a matching method called "Coarsened Exact Matching" (CEM) (Iacus et al. 2011, 2012), as it possesses a number of desirable statistical properties that other matching methods (e.g. propensity score matching) do not possess. First, CEM is a "Monotonic Imbalance Bounding" (MIB) estimator, which means that the maximum degree to which a variable can be out of balance is pre-specified. Second, CEM meets the "congruence principle," which states that the data space and the analysis space should be identical. Finally, and critically for our purposes, CEM automatically restricts the data to common support.

The CEM algorithm proceeds in four stages. First, we create a copy X^* of the covariates X. Second, we coarsen X^* by pre-specified cutoff point or binning algorithm. Third, for each unique observation of X^* we create one stratum and place the observations in those strata. Finally, we drop any observation that does not belong to a stratum containing at least one treated and one control unit, and then assign the remaining strata to the original data X (Blackwell et al. 2009). After pruning observations, we can then conduct the analysis by regressing Y on T and X, though when different numbers of treated and control units appear in different strata these regressions must be appropriately weighted (Iacus et al. 2012). Finally, it is important to note that the procedure can prune treated units thus changing the estimand to a "local average treatment on the treated."

Coarsening is central to CEM and, while our discrete covariates do not require further coarsening, it is necessary for the continuous covariates. Neither substantive knowledge of these continuous covariates nor inspection of the data reveals any obvious cutoff points. We thus use automated coarsening and employ two alternative approaches to assess the sensitivity of our results to the coarsening algorithm. Following Iacus et al. (2012), the first approach uses a leading binning rule developed by Shimazaki and Shinomoto (2007). As this approach imposes uniform bin sizes, our second approach uses the optimal k-means clustering algorithm described in Wang and Song (2011).[19] Both of these procedures yield coarsened data that results in too few matched units and, as such, we couple each approach with the inductive relaxation procedure described in Iacus et al. (2012). This relaxation procedure seeks to increase the number of matched units

---

[19]k-means clustering is a popular approach to cluster analysis. While standard k-means algorithms do not guarantee optimality, the Wang and Song (2011) algorithm is optimal for one-dimensional clustering. The number of clusters for each variable in this procedure is determined by the Bayesian information criterion.

by strategically reducing the number of bins or clusters for each variable. More specifically, the procedure discriminates between alternative relaxations by minimizing a multivariate imbalance measure.[20]

With the appropriate coarsening choices, we can then reasonably assume $E(\hat{\tau} - \tau) = 0$ (i.e. $\Delta_e = 0$ and our estimator is unbiased) and turn to discussing our empirical model.[21] The empirical model is as follows:

$$(4) \qquad\qquad y_{i,2012} = \alpha + \tau T_i + \beta X_{i,2005} + \varepsilon_i$$

where, for district $i = 1, 2, \ldots, N$ at time $t \in \{2005, 2012\}$, y denotes the dependent variable, T is the treatment indicator, X denotes a vector of independent variables (see Table 1 for a complete listing), and ε is the error term.[22] We estimate the model for all dependent variables (see Table 1) using the appropriate linear or non-linear model.[23] For each dependent variable, the model is estimated (1) without additional controls or matching; (2) with controls and without matching; (3) with controls and matching using Shimazaki-Shinomoto coarsening; and (4) with controls and matching using k-means coarsening. Finally, as the sphere of influence of the local governments is the province, all standard errors are clustered at the province level.

As will be seen, we find evidence that receipt of FOCAM transfers reduced municipal current expenditures (*current*). Following the discussion in Section 2, we hypothesize that municipalities with low absorptive capacity coped with the increased administrative burden of FOCAM transfers by reallocating administrative effort toward (away from) executing capital (current) expenditures. "An issue repeatedly highlighted in our interviews with municipal managers is the lack of necessary personnel . . . municipalities that enjoy a larger professional staff obtain higher execution rates" (Loayza et al. 2014, pg. 65). We thus proxy municipal absorptive capacity with (the log of) the number white-collar staff per 10,000 people (*staff*). To test our hypothesis, we simply interact *staff* with the treatment indicator in our model for current expenditures. We then estimate the model using OLS on three alternative samples: (1) without matching; (2) with matching using Shimazaki-Shinomoto coarsening; and (3) with matching using k-means coarsening.

---

[20]While Iacus et al. (2012) use a variable-by-variable approach to sequentially relax coarsening choices, we find that further gains can be made by considering combinations of relaxations. We thus randomly sample with 10,000 draws the entire space of feasible relaxations, though our solution concept remains the same as in Iacus et al. (2012). Computer code is available upon request.

[21]Importantly, we can assess the reduction in covariate imbalance by examining post-matching normalized differences (Imbens and Wooldridge 2009).

[22]To mitigate post-treatment bias, note that we only control for pre-treatment characteristics.

[23]That is, OLS is used for continuous outcome variables, logit models are used for binary outcomes, Poisson or negative binomial regressions are used for count data, and Tobit models are used for censored outcomes. The estimated treatment effect is then calculated as the marginal effect (evaluated at the mean) of T in each model.

5   Results

The results of our analysis are presented in Tables 2-4 and Figure 4. The final two columns of Table 2 present the post-matching normalized differences for all independent variables. The ND-SS (ND-KM) column presents normalized differences calculated after using CEM with the Shimazaki-Shinomoto (k-means) algorithm to coarsen continuous covariates. First, note the sample size (N). It is evident that the matching procedure reduces the number of observations from 1,492 to 1,095 and 1,107 for the Shimazaki-Shinomoto and k-means approaches, respectively. Second, relative to the pre-matching normalized differences (ND-PM), we see that the ND-SS and ND-KM columns present substantially reduced normalized differences. For example, the normalized difference for *literacy* is reduced from -0.29 to -0.02 and -0.03 in the ND-SS and ND-KM columns, respectively. Finally, for both approaches and across all covariates, we see that there remains no statistically significant differences in means after CEM.

CEM thus substantially reduces covariate imbalance and Table 3 presents the effect of matching on our estimates of treatment effect. Each row of Table 3 presents the impact estimates across alternative models for a given dependent variable. The second (baseline) column presents estimates from regressing a given dependent variable on the treatment indicator alone. The third (controls) column adds our independent variables to each regression (see Table 1 for the variable list). The final two columns (CEM-SS and CEM-KM) present our estimates when using CEM with the Shimazaki-Shinomoto and k-means coarsening, respectively. Recall that all estimates are marginal effects (evaluated at the mean) from the appropriate linear or non-linear model. That is, OLS is used for continuous outcome variables, logit models are used for binary outcomes, Poisson or negative binomial regressions are used for count data, and Tobit models are used for censored outcomes. Also recall that all standard errors are clustered at the province level.

Consider first the results for the (log of) municipal expenditures per capita variable (*expenditure*). While the baseline estimate shows that treated districts witnessed a statistically significant 19 percent increase in expenditures per capita, we see more subdued effects with the addition of control variables and matching. In particular, the CEM-SS and CEM-KM columns show estimated increases of 12 and 7 percent, respectively, but neither of these estimates is statistically significant. Recall, however, that FOCAM income must be directed to capital expenditures. Looking at the results for the (log of) municipal capital expenditures per capita variable (*capital*), we see relatively robust positive effects. Most notably, the CEM-SS and CEM-KM columns show statistically significant increases in *capital* of 22 and 17 percent, respectively. While treated districts thus witnessed increased capital expenditures, we conversely find evidence of a reduction in (the

log of) municipal current expenditures per capita (*current*). For example, our CEM-KM regression shows a statistically significant 11 percent reduction in *current*.

The reduction in current expenditures rationalizes the statistically insignificant estimates for overall expenditures, but it is an unexpected result. We further analyze this result below, but first examine the estimates associated with the other dependent variables. Given the estimated positive effect on capital expenditures, it is natural to ask whether there is an associated positive effect on local infrastructure. While estimates vary across alternative models, our preferred models (CEM-SS and CEM-KM) show robust positive impacts on *internet*, *construction*, *roads*, *sports*, and *cadastre*. Regarding *internet*, for example, our CEM-SS and CEM-KM results show that treated districts were, respectively, 8 and 11 percent more likely to have access to internet in 2012. To cite another example, the CEM-SS and CEM-KM regressions also show that treated districts constructed an additional 2.09 and 2.87 square kilometers of primary roads in 2012, respectively. The other infrastructure-related results can be interpreted analogously and we leave this exercise to the reader.

The final dependent variable in Table 3 to be discussed is the (log of) nighttime lights variable (*lights*). Given evidence of increased capital expenditure and an associated impact on local infrastructure, we might expect to see a positive impact on the nighttime lights variable. We, however, find only very limited evidence for this hypothesis. Our estimate with control variables suggests that treated districts witnessed a statistically significant 7 percent increase in *lights*. Our CEM-SS and CEM-KM results nevertheless show that this estimate is biased upward, and that more accurate point estimates are 2 and 4 percent, respectively. While we indeed find positive point estimates, neither the CEM-SS or CEM-KM results are statistically significant at any conventional level. Accordingly, increased capital expenditures and local infrastructure development do not appear to have translated into robust impacts on nighttime lights, though it is possible that additional impacts will emerge with the passage of more time.

We conclude this section by further examining the above result that FOCAM transfers reduced municipal current expenditures per capita (*current*). As stated, we hypothesize that this reduction in current expenditures is related to the absorptive capacity of municipal governments. That is, we hypothesize that municipalities with low absorptive capacity coped with the increased administrative burden of FOCAM transfers by reallocating administrative effort toward executing new capital expenditures. To test our hypothesis, we thus add to previous models a term that interacts *staff* with our treatment indicator. Table 4 presents results from this augmented model. Full regression results are presented for the model with all

controls and no matching, as well as for the CEM-SS and CEM-KM approaches. Recall that all regressions are estimated with OLS using standard errors clustered at the province level.

While the interaction term (*treated* × *staff*) is statistically insignificant in the regression without matching (second column of Table 4), the preferred CEM-SS and CEM-KM regressions show statistically significant point estimates of 0.12 and 0.13, respectively. To gain insight into the economic significance of these estimates, in Figure 4 we plot the heterogeneous treatment effects implied by the CEM-KM regression.[24] That is, we plot the point estimate and 95 percent confidence interval associated with the treatment effect on *current* as it varies with changes in *staff*. The results confirm our hypothesis: districts with low levels of absorptive capacity (i.e. *staff*) witness statistically significant reductions in current expenditures (i.e. *current*). More specifically, our results imply that districts with *staff* at one standard deviation below the mean witnessed an approximate 20 percent reduction in *current* (see Table 2 for descriptive statistics). Further, the point estimates also show that districts with high levels of *staff* witnessed increases in *current*, perhaps due to purchasing goods and services complementary to the new capital expenditures.[25]

6 Conclusions

We examined the economic effect of windfall gains by studying the impact of the Camisea Fund for Socioeconomic Development (FOCAM) in Peru. The rules governing FOCAM allocation created a natural experiment from which we were able to identify the effect of the transfers on municipal accounts, local infrastructure, and local development. Using a newly constructed district-level dataset for the years 2005 and 2012, we first found evidence of increased capital expenditures in treated districts. Second, we found that increased capital expenditures were associated with positive impacts on local infrastructure. That is, we found evidence of positive impacts on access to internet, licenses granted for new construction, and the building of primary roads, among other things. Finally, we found that districts with low absorptive capacity coped with the increased administrative burden of FOCAM transfers by reallocating administrative effort toward (away from) executing capital (current) expenditures.

This last results is particularly noteworthy, especially in the Peruvian context. The FOCAM transfer scheme is part of a larger fiscal reorganization effort in Peru that has attempted to use fiscal decentralization to reduce corruption and improve public service delivery. To this end, fiscal decentralization was accompanied

---

[24]A similar exercise could be undertaken with the CEM-SS model, but the results are virtually identical.

[25]Note, however, that zero always falls within the 95 percent confidence interval for the districts with high absorptive capacity. Also note that we tried a number of alternative specifications to permit some non-linearity in Figure 4. The associated hypothesis tests nevertheless revealed that the simple linear specification is most appropriate.

by stringent fiduciary requirements and rigid participatory budgeting guidelines for local governments and their elected officials. The anecdotal evidence suggests that many municipalities had difficulties complying with the drafted regulations and were unable to execute additional expenditures (see Loayza et al. [2014] and references therein for details). Our results, however, suggest that municipalities lacking absorptive capacity may not necessarily be unable to execute additional expenditures altogether, but may rather reallocate administrative effort to accommodate the additional expenditures. Thus, absent technical support and capacity-building efforts, such windfall gains may have unintended consequences.

While we believe that our results have important policy implications for improving the efficacy of windfall gains, our analysis has some limitations that should be acknowledged. First, data availability and compatibility issues precluded us from constructing a panel dataset including interim years. Second, also due to data availability issues, we were unable to analyze the potential consequences of reduced municipal current expenditures. Finally, while we were able to identify the average treatment on the treated, we lacked the appropriate instrumental variables to estimate the elasticity of our outcome variables to FOCAM transfers (i.e. the dose-response relationship). In future research we hope to extend our analysis by remedying some of these shortcomings, particularly by identifying the dose response.

References

Angrist, J. and Kugler, A. (2008). Rural windfall or a new resource curse? Coca, income, and civil conflict in Colombia. *The Review of Economics and Statistics*, 90(2):191–215.

Arndt, C., Jones, S., and Tarp, F. (2010). Aid, growth, and development: Have we come full circle? *Journal of Globalization and Development*, 1(2).

Arzaghi, M. and Henderson, J. (2005). Why countries are fiscally decentralizing. *Journal of Public Economics*, 89(7):1157–1189.

Bazzi, S. and Clemens, M. (2013). Blunt instruments: Avoiding common pitfalls in identifying the causes of economic growth. *American Economic Journal: Macroeconomics*, 5(2):152–186.

Becker, S., Egger, P., and von Ehrlich, M. (2013). Absorptive capacity and the growth and investment effects of regional transfers: A regression discontinuity design with heterogeneous treatment effects. *American Economic Journal: Economic Policy*, 5(4):29–77.

Bird, R. (1993). Threading the fiscal labyrinth: Some issues in fiscal decentralization. *National Tax Journal*, 46(2):207–227.

Bird, R. and Smart, M. (2002). Intergovernmental fiscal transfers: International lessons for developing countries. *World Development*, 30(6):899–912.

Blackwell, M., Iacus, S., King, G., and Porro, G. (2009). cem: Coarsened exact matching in Stata. *Stata Journal*, 9(4):524–546.

Brollo, F., Nannicini, T., Perotti, R., and Tabellini, G. (2010). The political resource curse. NBER Working Paper No. 15705.

Brunnschweiler, C. and Bulte, E. (2008). The resource curse revisited and revised: A tale of paradoxes and red herrings. *Journal of Environmental Economics and Management*, 55(3):248–264.

Burnside, C. and Dollar, D. (2000). Aid, policies, and growth. *The American Economic Review*, 90(4):847–868.

Caselli, F. and Michaels, G. (2009). Do oil windfalls improve living standards? Evidence from Brazil. NBER Working Paper No. 15550.

Chen, X. and Nordhaus, W. (2011). Using luminosity data as a proxy for economic statistics. *Proceedings of the National Academy of Sciences*, 108(21):8589–8594.

Contraloría General de la República del Perú (2014). *Estudio del proceso de descentralización en el Perú*. Contraloría General de la República, Lima, Peru.

Dalgaard, C. and Olsson, O. (2008). Windfall gains, political economy and economic development. *Journal of African Economies*, 17(suppl 1):72–109.

De Mello, L. (2000). Fiscal decentralization and intergovernmental fiscal relations: A cross-country analysis. *World Development*, 28(2):365–380.

Dillinger, W. (1994). Decentralization and it implications for urban service delivery. Urban Management and Municipal Finance Discussion Paper No. 16, The World Bank.

Dube, O. and Vargas, J. (2013). Commodity price shocks and civil conflict: Evidence from Colombia. *The Review of Economic Studies*, 80(4):1384–1421.

Gadenne, L. and Singhal, M. (2013). Decentralization in developing economies. NBER Working Paper No. 19402.

Hansen, H. and Tarp, F. (2001). Aid and growth regressions. *Journal of Development Economics*, 64(2):547–570.

Henderson, J., Storeygard, A., and Weil, D. (2012). Measuring economic growth from outer space. *The American Economic Review*, 102(2):994.

Holland, P. (1986). Statistics and causal inference. *Journal of the American Statistical Association*, 81(396):945–960.

Iacus, S., King, G., and Porro, G. (2011). Multivariate matching methods that are monotonic imbalance bounding. *Journal of the American Statistical Association*, 106(493):345–361.

Iacus, S., King, G., and Porro, G. (2012). Causal inference without balance checking: Coarsened Exact Matching. *Political Analysis*, 20(1):1–24.

Imbens, G. and Wooldridge, J. (2009). Recent developments in the econometrics of program evaluation. *Journal of Economic Literature*, 47(1):5–86.

Inter-American Development Bank (2002). Program for institutional strengthening and environmental and social management support for the Camisea Gas Project: Loan proposal. Retrieved from www.iadb.org/en/projects/project-description-title,1303.html?id=PE0233. Accessed September 23, 2015.

Inter-American Development Bank (2004). Informe resumen de la ejecución de los compromisos ambientales y sociales del Projecto Camisea. Retrieved from idbdocs.iadb.org/wsdocs/getdocument.aspx?docnum=455833. Accessed September 23, 2015.

King, G. and Zeng, L. (2006). The dangers of extreme counterfactuals. *Political Analysis*, 14(2):131–159.

Loayza, N., Rigolini, J., and Calvo-González, O. (2014). More than you can handle: Decentralization and spending ability of Peruvian municipalities. *Economics & Politics*, 26(1):56–78.

Meyer, B. (1995). Natural and quasi-experiments in economics. *Journal of business & economic statistics*, 13(2):151–161.

Michalopoulos, S. and Papaioannou, E. (2013). Pre-colonial ethnic institutions and contemporary african development. *Econometrica*, 81(1):113–152.

Ministerio de Economía y Finanzas (2015). Fondo de Desarrollo Socioeconómico de Camisea - FOCAM. Retrieved from www.mef.gob.pe/index.php?option=com_content&view=article&id=451&Itemid=100962&lang=es. Accessed September 23, 2015.

Munilla, I. (2010). *People, power, and pipelines: Lessons from Peru in the governance of gas production revenues*. Oxfam America, Boston, MA.

Paler, L. (2011). The subnational resource curse: Causes, consequences, and prescriptions. Revenue Watch Institute.

Rajan, R. and Subramanian, A. (2008). Aid and growth: What does the cross-country evidence really show? *The Review of Economics and Statistics*, 90(4):643–665.

Sachs, J. and Warner, A. (1995). Natural resource abundance and economic growth. NBER Working Paper No. 5398.

Shimazaki, H. and Shinomoto, S. (2007). A method for selecting the bin size of a time histogram. *Neural Computation*, 19(6):1503–1527.

The Economist (2003). Gas for peru v green imperialism. Retrieved from www.economist.com/node/1978816. Accessed September 23, 2015.

van der Ploeg, F. (2011). Natural resources: Curse or blessing? *Journal of Economic Literature*, 49(2):366–420.

van der Ploeg, F. and Poelhekke, S. (2010). The pungent smell of "red herrings": Subsoil assets, rents, volatility and the resource curse. *Journal of Environmental Economics and Management*, 60(1):44–55.

Vicente, P. (2010). Does oil corrupt? Evidence from a natural experiment in West Africa. *Journal of Development Economics*, 92(1):28–38.

Walsh Peru S.A. (2002). EIA Gasoducto Camisea - Lima. Retrieved from www.southperupanel.org/files/eias/03.%20CAMISEA-LIMA/VOL%20I/7.0%20Alternativas.pdf. Accessed September 23, 2015.

Wang, H. and Song, M. (2011). Ckmeans.1d.dp: Optimal k-means clustering in one dimension by dynamic programming. *The R Journal*, 3(2):29–33.
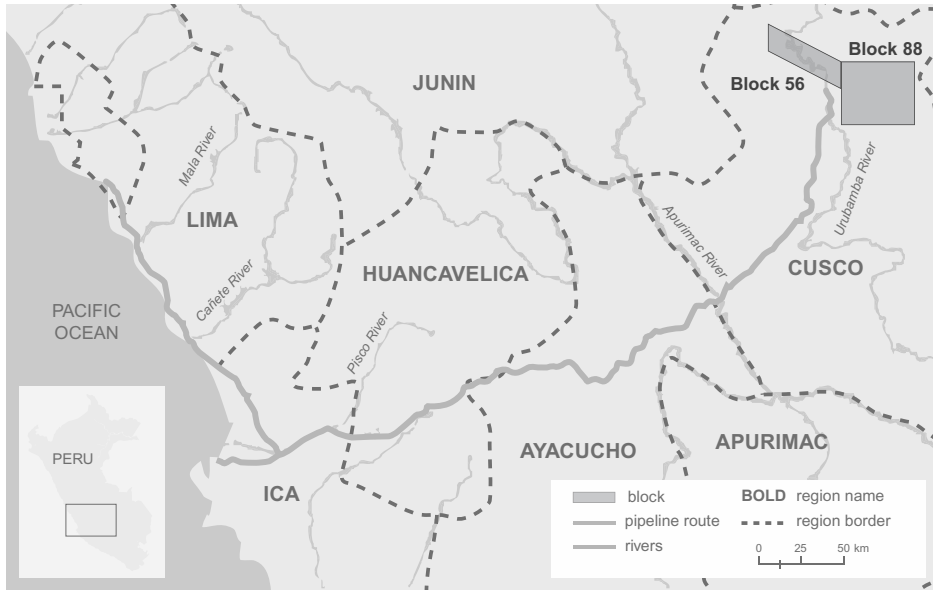
Appendix

Figure 1: The Camisea Gas Project (*Source*: Munilla [2010])
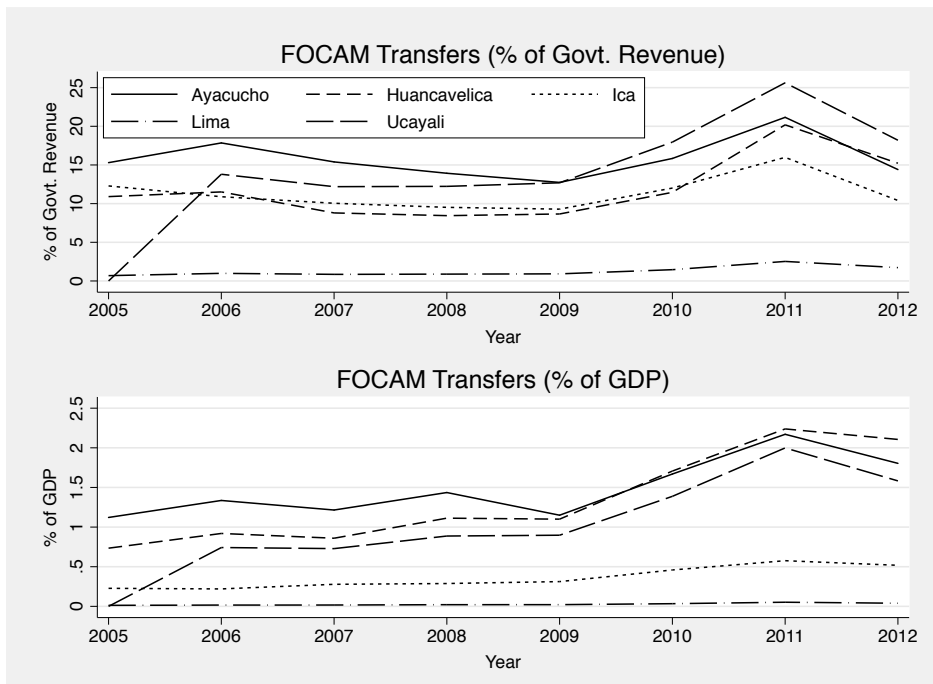

Figure 2: FOCAM transfers as a percentage of government revenue and GDP (2005-2012)
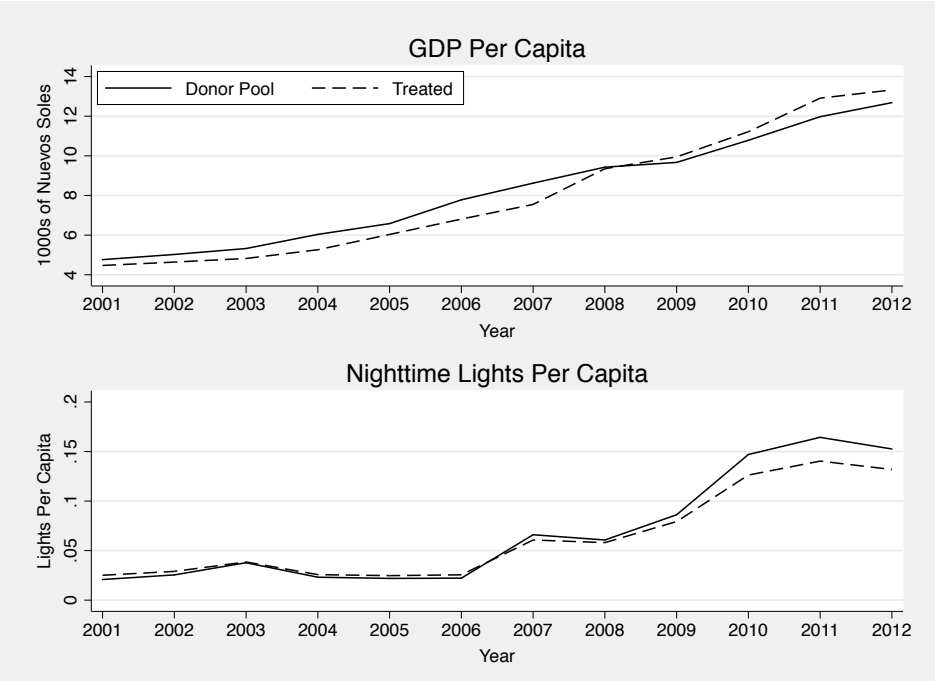
Figure 3: GDP and nighttime lights per capita for treated and donor pool regions (2001-2012)

Table 1: Independent and dependent variable definitions

Variable	Definition
	Independent Variables
lights	log of annual nighttime lights per 10,000 people
highlands	district located in highlands region
coast	district located in coastal region
area	log of district land area (square kilometers)
population	log of district population size
dependents	district dependency ratio
literacy	district literacy rate
transfers	log of non-FOCAM transfer income per 10,000 people[a]
staff	log of number of white-collar municipal staff per 10,000 people
	Dependent Variables
expenditure	log of municipal expenditures per capita (current *nuevos soles*)
capital	log of municipal capital expenditures per capita (current *nuevos soles*)
current	log of municipal current expenditures per capita (current *nuevos soles*)
internet	district has access to internet
business	number of licenses granted for new businesses
construction	number of licenses granted for new construction (e.g. houses, businesses, etc.)
roads	primary roads constructed (square kilometers)
sports	number of stadiums
culture	number of libraries
health	number of health-related facilities (i.e. clinics, health centers, etc.)
water	water receives treatment
cadastre	district has cadastre
lights	log of annual nighttime lights per 10,000 people

[a] The non-FOCAM transfer variable is in 10,000 current *nuevos soles*. It is calculated as the sum of *Canon Minero*, *Canon* and *Sobrecanon Petrolero*, *Canon Forestal*, *Canon Gasífero*, *Canon Hidroenergético*, *Canon Pesquero*, *regalias mineras*, and FONCOMUN transfers.

Table 2: Variable means by treatment status and normalized differences (2005)[a]

Variable	Treated	Donor Pool	ND-PM	ND-SS	ND-KM
lights	3.65 (2.08)	4.11 (2.15)	-0.15	-0.02	0.04
highlands	0.83 (0.38)	0.63 (0.48)	0.31**	0.20	0.13
coast	0.16 (0.37)	0.18 (0.39)	-0.04	0.00	0.00
area	5.26 (1.15)	5.41 (1.45)	-0.08	-0.08	0.00
population	8.12 (1.14)	8.54 (1.23)	-0.26**	-0.09	-0.09
dependents	83.41 (15.54)	74.73 (14.96)	0.40***	0.20	0.07
literacy	0.78 (0.08)	0.81 (0.07)	-0.29**	-0.02	-0.03
transfers	5.14 (1.01)	4.98 (1.27)	0.10	0.01	0.07
staff	3.31 (0.71)	3.13 (0.69)	0.18*	0.14	0.04
N	246	1,246	1,492	1,095	1,107

[a] Standard deviations for treated and donor pool variable means are presented in parentheses. P-values for normalized differences are from OLS regressions of the independent variable on the treatment indicator. Standard errors in those regressions are clustered at the province level. P-values <0.01, 0.05, and 0.10 correspond to ***, **, and *, respectively.

Table 3: Estimates of treatment effect for alternative models and dependent variables[a]

Variable	Baseline	Controls	CEM-SS	CEM-KM
expenditure	0.19**(0.08)	0.04 (0.07)	0.12 (0.08)	0.07 (0.08)
capital	0.31***(0.09)	0.13 (0.09)	0.22**(0.10)	0.17*(0.10)
current	-0.03 (0.08)	-0.13**(0.06)	-0.09 (0.07)	-0.11*(0.07)
internet	-0.01 (0.05)	0.05*(0.03)	0.08*(0.05)	0.11**(0.04)
business	-15.29**(6.78)	4.27*(2.61)	3.44*(1.89)	1.73 (1.98)
construction	39.30 (27.47)	7.90***(2.36)	3.72***(1.37)	5.91***(1.95)
roads	2.46 (1.79)	5.18***(1.77)	2.09***(0.56)	2.87***(0.89)
sports	0.19**(0.09)	0.24***(0.08)	0.15*(0.08)	0.21**(0.09)
culture	-0.12 (0.09)	-0.04 (0.07)	-0.01 (0.07)	-0.01 (0.06)
health	-0.04 (0.08)	0.09 (0.09)	0.04 (0.09)	0.02 (0.07)
water	0.08**(0.04)	0.07*(0.04)	0.08**(0.04)	0.06 (0.04)
cadastre	0.06 (0.04)	0.11**(0.05)	0.12**(0.06)	0.10*(0.06)
lights	0.20 (0.20)	0.07*(0.04)	0.02 (0.04)	0.04 (0.04)
N	1,492	1,492	1,095	1,107

[a] All estimates of treatment effect are marginal effects (evaluated at the mean) from the appropriate linear or non-linear regression model. Standard errors (in parentheses) are clustered at the province level. P-values <0.01, 0.05, and 0.10 correspond to ***, **, and *, respectively.

Table 4: Triple difference estimates for the log of municipal current expenditures per capita[a]

Variable	Controls	CEM-SS	CEM-KM
treated	-0.26*(0.14)	-0.50***(0.17)	-0.54***(0.21)
lights	0.02**(0.01)	0.02**(0.01)	0.05***(0.01)
highland	0.15***(0.05)	0.08 (0.06)	0.20**(0.09)
coast	0.15**(0.07)	0.16*(0.08)	0.03 (0.11)
area	0.08***(0.02)	0.05***(0.02)	0.07***(0.03)
population	-0.07***(0.02)	-0.09***(0.02)	-0.10***(0.02)
dependents	-0.003 (0.003)	-0.0002 (0.003)	0.002 (0.003)
literacy	-0.16 (0.33)	0.26 (0.38)	0.96**(0.42)
transfers	0.11***(0.02)	0.13***(0.04)	0.06**(0.02)
staff	0.36***(0.03)	0.23***(0.04)	0.23***(0.05)
treated × staff	0.04 (0.04)	0.12**(0.05)	0.13**(0.06)
R^2	0.43	0.38	0.39
N	1,492	1,095	1,107

[a] All regressions are estimated with OLS. Standard errors (in parentheses) are clustered at the province level. P-values <0.01, 0.05, and 0.10 correspond to ***, **, and *, respectively.
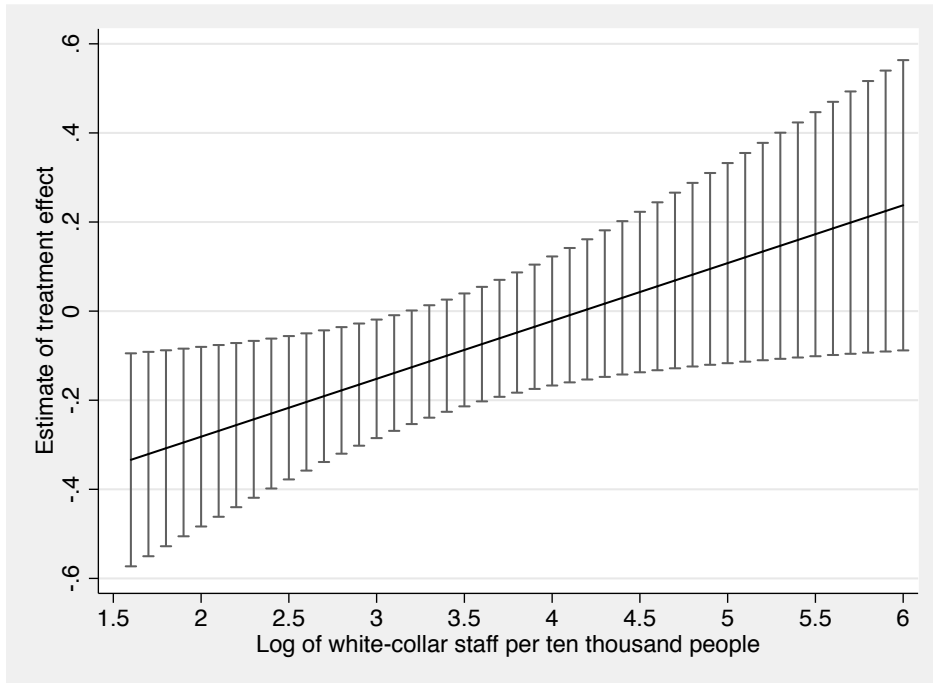
Figure 4: Heterogeneity of impact on the log of municipal current expenditures per capita