**Online Appendix for: Do Oil Windfalls Improve Living Standards? Evidence from Brazil**

Francesco Caselli and Guy Michaels

## Unpublished Appendix 1: Rules governing the allocation of royalties from oil and gas in Brazil

This appendix is based on ANP (2001): Guia dos Royalties do Petróleo e do Gás Natural 2001.

The current allocation of royalties is the result of a series of incremental legislative changes between 1953 and 1998. The incremental nature of the legislation has resulted in a rather complicated structure, which we now try to describe.

The total amount of royalty payments from each oilfield is the sum of two components. The first component is a fixed 5% of the value of the oil extracted. We call it the fixed quota. The second is a further percentage that must be between 0 and 5%. We call this the variable quota. However, even the variable quota is almost always set at the maximum of 5%. This is because the legislation authorizes ANP to assign a quota less than the maximum only in the case of lower-quality or higher-risk fields. As a result about 90% of the oilfields and all of the large oilfields pay between 9.1 and 10% in royalties. Another 9% pays between 8.1 and 9%. Only 1% pays less than 8%. The weighted average royalty is 9.8%.

This does not mean that the distinction between fixed quota and variable quota is irrelevant, though, because the sets of recipients of the two quotas, and the way the quotas are distributed among the various recipients, are very different. In particular, the fixed quota is divided as follows. For onshore fields, 70% to "producing" states, 20% to "producing" municipalities, and 10% to municipalities with significant offshore-oil related infrastructure (essentially, terminals for bringing offshore oil to land). For offshore fields, 30% to "facing" states, 30% to "facing" municipalities, 20% to the Navy, 10% to a "special fund" to be divided between all states and all municipalities, and 10% to municipalities with significant oil-related infrastructure. We come back to the definitions of "producing" and "facing" below.

The variable quota has even more recipients. For onshore oilfields, 52.5% goes to "producing" states, 25% to the Ministry of Science, 15% to "producing" municipalities, and 7.5% to municipalities "affected" by operations connected with the landing of offshore oil. For offshore oil, 22.5% to the Ministry of science, 22.5% to "facing" states, 22.5% to "facing" municipalities, 15% to the Navy, 7.5% to the "special fund," and 10% to "affected" municipalities

If one were to combine the percentages from the fixed and the variable quota, in the (typical) case of an oilfield paying the maximum royalty (i.e. 10%) then the total share going to *producing* municipalities in an onshore oilfield is 18%. For *offshore* fields, the percent of a 10% royalty going to *facing* municipalities would be 26%. Unfortunately, however, in the case of offshore oil things are not so simple, because the definition of "facing" is different in the case of the fixed quota and in the case of the variable quota. We thus now turn to the definitions of producing and facing.

For onshore oil, a state (municipality) is a "producing" state (municipality) vis-à-vis a certain oilfield if and only if there are wells tapping into that particular oilfield that are inside the state's (municipality's) borders. Each producing state (municipality) participates into the total royalties allocated to producing states (municipalities) from a certain oilfield in proportion to the *output share* of the wells situated in that state (municipality) in the total oilfield output. This is true both for the fixed and for the variable quota.

For offshore oil, things get quite complicated. First of all, state and municipality maritime boundaries are needed. The relevant legislation assigns this task to the Brazilian Geographical Institute (BGE), and this has resulted into two distinct sets of borders. The first set is based on *perpendicular lines*. It begins by picking 25 points on the Brazilian coast, and connecting them by straight lines. It is necessary to "linearize" the coast because the fractal nature of the coastline would otherwise make it impossible to draw perpendicular lines going out to sea. The 25 points include all the points at which two coastal state's boundaries reach the shore, but they also include a few extra points to accommodate extreme irregularities of the coastline inside a state.

Once the coastline has been 'linearized,' parallel lines going out to see are drawn from the points where the state borders reach the coastline. These lines are deemed to be the continuation of the state border onto the continental shelf, and they "end" when they meet the outside boundary of the Brazilian continental shelf (i.e. the end of Brazil's territorial water).

Municipality boundaries based on perpendicular lines follow similar principles, with a few adjustments. First, in the states of Rio De Janerio and Sao Paulo, a few more points are added to the 'linearization' of the coast so the linearized version used for municipal boundaries is a bit more jagged than the one used for state ones. Second, municipal borders end either when they reach the continental shelf boundaries (as was the case for state borders), or when they reach the state boundary.

The second set of boundaries is based on *parallel lines*. It amounts to identifying state (municipal) boundaries with the parallel passing by the point where state (municipal) boundaries reach the coast (and ending at the continental-shelf boundary).

Once state boundaries are available, a municipality is "facing" a certain oilfield if there are wells tapping into this oilfield that lie inside the municipality's maritime border. If the two sets of borders map the same well into two different municipalities they both have equal rights to the royalties.

This is far from the end of the story, though. For, each identified facing municipality must share the "facing" quota with a set of neighboring municipalities, called the geo-economic area. The construction of the geo-economic area begins by identifying the "geographic mesoregion" to which the facing municipality belongs. The geographic mesoregion is a purely geographic construct that exists independently of the royalty allocation mechanism. Each municipality belongs to one and only one geographic mesoregion.

Next, within the mesoregion, IBGE identifies a "main production zone." The facing municipality must always belong to the main production zone. In addition, this zone includes municipalities with at least three of the following: (i) infrastructure for processing, treating, storing, and shipping oil (excluding pipelines); (ii) infrastructure supporting exploration, extraction, and shipment of oil (ports, airports, manufacture and maintenance of oil-rig equipment, etc.). However, it turns out that very few (i.e. eight) municipalities fulfill this criterion, so the vast majority of the municipalities in the main production zone are the municipalities facing the oil wells.

Once the main production zone is identified, the geo-economic area is the union of two sets: all the municipalities in the mesoregion, and all the municipalities that border the main production zone. In several cases, of course, the latter set is a subset of the former, but often that is not the case. Municipalities in a geo-economic area that are not in the main production zone are assigned to the "geographic zone contiguous to the main production zone." Municipalities in the geo-economic area whose territory is crossed by pipelines transporting offshore oil/gas are assigned to the "secondary production zone." There are only 8 of these. All municipalities in the secondary production zone therefore also belong to the geographically contiguous zone, but if one municipality receives royalties by virtue of being in the secondary zone then it is excluded from division of royalties based on the contiguity criterion.

Given the total royalties from the fixed quota going to a certain geo-economic area based on the "facing" principle (i.e. 30% of 5%), there is a first round of allocation that works as follows: 60% to the main zone, 10% to the secondary zone, and 30% to the contiguous zone. Next, within each zone, each municipality's share depends on its population size.[1] Recall that in practice, the main zone tends to be constituted almost exclusively by the facing municipality, so in practice we should expect close to 60% of the 30% of the 5% to go to facing municipalities.

Note that, as we have seen, the same oil well may well be inside two municipality's borders, depending on the perpendicular or parallel principle. If the two municipalities are in the same geo-economic area this fact has no implications whatsoever for the allocation of royalties, as the identity of the municipality in whose border the well lies does not affect the allocation within the geo-economic area. But, of course, if the two municipalities are in different geo-economic areas then the number of municipalities sharing in the royalties increases accordingly.

Finally, how is the total "facing" component of the royalty allocated between geo-economic areas? The principle is the same as for onshore oil, i.e. each geo-economic area receives royalties in proportion to the *output share* of the wells situated inside its maritime borders in the total oilfield output.

Things are very different, and much simpler, for the variable quota. First, here it is facing municipalities only: nothing goes to the geo-economic areas. (Although there is a 7.5% separate

[1] However, if one municipality in the main zone has at least three pieces of land-based infrastructure for processing, treatment, storage, and disposal of offshore oil it must receive at least one-third of the overall royalties going to the zone in which it sits. Hence, if the allocation based on population implies that this municipality receives less than one-third of the total, a new allocation is made where it receives one-third and the remainder is divided among the others based on the population criterion. In practice, only two municipalities satisfy this criterion.

quota for municipalities with infrastructure, but this is outside the “facing” quota). Second, and more importantly, the identification of a facing municipality is no longer based on the location of wells, but on the location of fields. In particular, for each field, the set of facing municipalities is the set of municipalities whose borders’ extensions on the continental shelf (whether drawn with perpendiculars or parallels) contain any portion of the field.

The allocation of the overall 22.5% of the variable quota among facing municipalities is pro-rated based on the simple average of the municipality’s share in the total field area based on perpendiculars and based on parallels. Hence, if a1(m,f)% of field f lies inside the maritime borders of municipality m by the perpendicular-line criterion, and a2(m,f)% according to the parallel-line criterion, then the royalties of municipality m (based on the “facing” criterion alone, and only on the variable quota) are

$$(1/2) \times [a1(m,f) + a2(m,f)] \times 0.225 \times q(f),$$

where q(f) is the value of the output of field f. Accordingly, the formula by which we seek to assign offshore output to municipalities is

$$(1/2) \times [a1(m,f) + a2(m,f)] \times q(f).$$

**Unpublished Appendix 2: Data sources, construction, definitions, and limitations**

A2.A From Municipalities to AMCs

IPEA has multiple AMC partitions, for different spans of time (e.g. 1920-2000, 1940-2000, etc.). Because of the progressive fragmentation of municipalities, the longer the time span the coarser the disaggregation (i.e. the fewer the AMCs). The earliest data we use in this paper are for 1970, so we use 1970-2000 AMCs.

The crosswalk from municipalities to AMCs we use maps municipalities that existed in 1997 into 1970-2000 AMCs. One slight complication is that the municipal-level variables we use are for 2000-2005, so there are a few municipalities that did not exist in 1997. In these few cases before applying the crosswalk we first assigned the new municipalities to 1997 municipalities "by hand."

We obtained the cross-walk from Eustaquio Reis of IPEA, to whom we are grateful. At the time of writing the cross-walks were also available online, e.g. at http://www.timthomas.net/.

A2.B Additional information on the construction of AMC oil output

For each producing oilfield and for every month since August 1998 ANP reports the reference price used to calculate royalties from oil and gas. For the same period, ANP also lists the quantity of oil and gas produced in each oilfield in each month. The value of annual output is the sum over the year of the values of monthly outputs, obtained by multiplying the reference price by the quantity produced.

The reference price is the maximum between the actual sale price of the oil extracted in a particular field and an imputed sale price (for oil delivered to Petrobras-owned refineries) based on prevailing world-market prices for oil with similar chemical composition. In practice, the reference price is essentially indistinguishable from the market price, so our measure of field-specific oil revenues should be very accurate. For details on the reference price see ANP (2001).

At the time of writing the reference prices by month and field were at http://www.anp.gov.br/?id=534, and the quantities extracted were at http://www.anp.gov.br/?id=532.

The value of offshore oil is allocated to municipalities using the Percentuais Médios de Confrontação for February 2008. We have done some checks to make sure that these shares do indeed reflect the stated geographical principles that should form the basis for these percentages. In most cases, they seemed fairly consistent. However, there were a few smaller oilfields for which the allocation of percentages did not seem consistent with the stated criteria. We have been unable to establish what alternative criteria had been used in these cases. We have also checked the shares published for other months (e.g. March 2009) and, as expected, found little variation over time.

At the time of writing, a ling to the most recent Percentuais was at http://www.anp.gov.br/?pg=9188&m=&t1=&t2=&t3=&t4=&ar=&ps=&cachebust=1258621993687.

In order to identify onshore oilfields, and then allocate their output among municipalities, we first downloaded from Banco de Dados de Exploração e Produção (BDEP), a database maintained by ANP, a list of all fields that had passed the stages of development and were already producing in December 2007, as well as information on their geographical location. We then defined a field as onshore if no Confrontação shares existed for this field. For these onshore fields, we compared their geographical contours with those of the boundaries of Brazilian municipalities, obtained from the Instituto Brasileiro de Geografia e Estatistica, IBGE. This allowed us to establish the geographical relationship between the various oilfields and the various municipalities.

At the time of writing the BDEP map database was at http://maps.bdep.gov.br/website/maps/viewer.htm. The date of our download was March 3, 2008. The IBGE map database was at http://mapas.ibge.gov.br/divisao/viewer.htm. For reasons discussed above we use the 1997 boundaries.

The criteria for defining oil municipalities, as well as offshore and onshore oil municipalities are based on the same criteria for the classification of oilfields above. Namely, onshore oil municipalities are those that lie above an onshore oilfield, while offshore oil municipalities are those with a positive Confrontação share.

A2.C Variables downloaded from IPEA

Any variable used in the paper whose source is not explicitly given in this appendix or in the text should be understood to have been directly downloaded (at the AMC 1970-2000 level) from the IPEA data web site, ipeadata. Most variables' definitions are self-explanatory and/or given in the text of the paper. Here we only give a few additional details on a few that merit comment.

*Population.* In estimating (1) and (2) we typically normalize both left- and right-side variables by population. Population data up to the year 2000 comes from IPEA, based on the Brazilian Censuses. To calculate population for years after 2000 we inflated the 2000 population from the Census estimates of the percentage change in population residing in each AMC on 1 July of each year, as reported by IPEA's. Similarly, there is one instance where we need population data for 1992, and, again, we used a similar interpolation from the 1991 Census.

At the time of writing, the URL for ipeadata was http://www.ipeadata.gov.br/

A2.D Variables not downloaded from IPEA

*Royalties*. The amount of oil royalties received by each municipality in each year is available from ANP. At the time of writing the URL for the royalty data was http://www.anp.gov.br/?pg=9080&m=&t1=&t2=&t3=&t4=&ar=&ps=&cachebust=1258622130031.

*Coastal Dummy*. While most of the geographic controls (latitude, longitude, distance from state and federal capital, state-capital dummy, and area) come from IPEA, we created an indicator for whether a municipality is adjacent to the coast using the IBGE GIS data referred to above (an AMC is coastal if at least one of its municipalities is coastal).

*Rooms at home.* The number of rooms at home is a variable we computed ourselves from Censo Demográfico (question V0203). The other variables in the first 8 columns of Table 6 are also constructed from census data but are directly available from IPEA at the AMC level.

At the time of writing the URL for Censo Demográfico was http://www.ibge.gov.br/lojavirtual/fichatecnica.php?codigoproduto=7791&midia= and http://www.ibge.gov.br/lojavirtual/fichatecnica.php?codigoproduto=7792&midia=

*Municipal roads.* We downloaded the raw data for the construction of our road variables from the Ministry of Transport's web page. The dataset lists approximately 11,200 federal, state, and municipal paved roads, and provides GIS shapefiles. Combining these with the shapefiles for our AMCs we were able to determine the list of AMCs through which each road passes. The dataset also gives the extension of each road. We then apportioned the overall extension of each road equally among the AMCs that "share" in that road.

The combined extension of the paved roads in the dataset is approximately 156,000km. According to Wikipedia the extension of paved roads in Brazil is approximately 200,000km, so it is possible our data are not exhaustive (http://en.wikipedia.org/wiki/List_of_Brazilian_Highways). Nevertheless, at least for paved roads there is no reason that we are aware of to expect selection to specifically affect oil AMCs. (The coverage of our data is much poorer for unpaved roads, which according to Wikipedia account for the vast majority of Brazilian roads, but are only a small fraction of those in our dataset).

We should also note that in our sample there are only approximately 500 municipal (paved) roads, so most municipalities are simply not in the business of building and managing roads. Nevertheless, it is interesting to see whether oil money induces some to enter this business.

At the time of writing the path to downloading the road data was the following. From http://www.transportes.gov.br follow links to BASE DE DADOS GEORREFERENCIADA, within which the relevant files can be found in the folder PORTFOLIOS DE PROJETOS\PORTFOLIOS\Modais.

*Teachers and Classrooms.* We constructed the education data inputs using the microdata from the school census (Censo Escolar) an annual comprehensive collection of administrative information on all Brazilian schools, carried out by INEP. The cleanest and most up to date version of this data is widely available among researchers in the field.

*Clinics and Hospitals.* The health data come from Pesquisa Assistência Médico Sanitária (AMS) 2002, a comprehensive survey of the Brazilian health system carried out by the Ministry of

Health. Hospitals are identified as “estabelecimentos” with “atendimento com internação,” while clinics are “estabelecimentos” with “atendimento sem internação.”

At the time of writing data on the Pesquisa AMS 2002 was available for downloaded at http://tabnet.datasus.gov.br/cgi/deftohtm.exe?ams/cnv/namsabr.def.

*Transfers received by households.* Computed by us from Censo Demográfico (URL above).

*Household income*. Computed by us (see text) from Censo Demográfico (URL above). We refer as household income per capita to the following construct. (i) For each individual record in the census divide household income by household size; (ii) sum all the resulting individual incomes (using census weights); (iii) divide by population. It may be more accurate to refer this to as a measure of personal income, rather than household income, but we use household income to flag the fact that the raw data report household income for each individual.

*Police operations.* These are official administrative data available from the Federal Police’s web page: http://www.dpf.gov.br/DCS/Resumo_OP_200X.html, where X=5, 6, 7, 8 or 3-2004 (for example, for 2008 the link is here: http://www.dpf.gov.br/DCS/Resumo_OP_2008.html; 2003 and 2004 are reported together).

A2.E Deflation

Many of the economic variables we used were obtained directly in R$2000. Other variables were denominated in nominal R$, and we converted them to R$2000 using a CPI index from IPEA (Índice Nacional de Preços ao Consumidor - INPC).

**Unpublished Appendix 3: Regressions underlying estimates of cumulative oil-related revenues over the 1990s.**

As explained in the paper, in order to estimate the cumulative revenues flow associated with oil production in 2000 we run a set of regressions of total municipal revenues in each preceding year on oil output in 2000. The cumulative effect of one Real of oil in 2000 on revenues between 1991 and 2000 is in principle the sum of all these coefficients. The results from such an exercise are reported in the following table.

| Year | AMCs | fraction reporting | coefficient | s.e. |
|---|---|---|---|---|
| 1991 | 3,475 | 0.97 | 0.0019 | 0.0008 |
| 1992 | 3,522 | 0.98 | 0.0013 | 0.0005 |
| 1993 | 3,512 | 0.98 | 0.0026 | 0.0005 |
| 1994 | 3,465 | 0.97 | 0.0072 | 0.0014 |
| 1995 | 3,453 | 0.96 | 0.0098 | 0.0024 |
| 1996 | 3,212 | 0.90 | 0.0038 | 0.0021 |
| 1997 | 3,422 | 0.95 | 0.0025 | 0.0022 |
| 1998 | 2,689 | 0.75 | -0.0032 | 0.0036 |
| 1999 | 2,705 | 0.75 | 0.0019 | 0.0044 |
| 2000 | 3,484 | 0.97 | 0.0311 | 0.0039 |