BANCO CENTRAL DE RESERVA DEL PERÚ

The Blessing of Natural Resources: Evidence from a Peruvian Gold Mine

Fernando M. Aragón * Juan Pablo Rud **

* Department of Economics, London School of Economics and Political Science
** Department of Economics, Royal Holloway, University of London

DT. N° 2009-015
Serie de Documentos de Trabajo
Working Paper series
Diciembre 2009

The Blessing of Natural Resources: Evidence from a Peruvian Gold Mine*

Fernando M. Aragón[†] Juan Pablo Rud[‡]

November 2009

Abstract

This paper studies the impact of Yanacocha, a large gold mine in Peru, on the local population. Using annual household data from 1997 to 2006, we find robust evidence of a positive effect of the mine's demand of local inputs on real income. The effect, an average income increase of 1.7% per 10% additional mine's purchases, is only present in the mine's supply market and surrounding areas. We also find evidence of improvements on measures of welfare and reduction of poverty. We examine and rule out that our results are driven by increased public expenditure associated to the mining revenue windfall. Using a spatial general equilibrium model, we interpret these results as evidence of net welfare gains generated by the mine's backward linkages and its multiplier effect.

Keywords: Natural resources, mining, local development.

JEL classification: O13, O18, Q32, Q33, R20

*We are particularly grateful to Tim Besley for his support and advice during this research project. We also thank Nava Ashraf, Oriana Bandiera, Konrad Burchardi, Francesco Caselli, Dave Donaldson, Claudio Ferraz, Greg Fischer, Maitreesh Ghatak, Guy Michaels, Gerard Padró i Miquel, Torsten Persson and seminar participants at LSE, Royal Holloway, the European Winter Meeting of the Econometric Society and Oxcarre for useful comments and suggestions.

[†]Department of Economics, London School of Economics and Political Science, Houghton Street, London WC2A 2AE, United Kingdom; Tel: +44 (0)20 7852 3539; Fax: +44-(0)20-7955-6951; Email: f.m.aragon@lse.ac.uk

[‡]Department of Economics, Royal Holloway, University of London, Egham, Surrey, TW20 0EX, United Kingdom; Tel: +44 (0)1784 27 6392; Email: juan.rud@rhul.ac.uk

> The grudge against what has become known as the "enclave" type of development is due to this ability of primary products from mines, wells, and plantations to slip out of a country without leaving much of a trace in the rest of the economy.
>
> Hirschman (1958, p. 110)

1 Introduction

Access to natural resources has hardly been associated with development. A body of theoretical and empirical literature suggests that natural resources may not translate into better living standards or may even hinder development (Sachs and Warner, 1999; Sachs and Warner, 2001; Caselli and Michaels, 2009). The current explanations for this "natural resource curse" focus on conflict, bad institutions, ineffective government or displacement of other productive sectors like in the Dutch disease (Benjamin et al., 1989; Torvik, 2002; Caselli, 2006). However, the available evidence tell us little about alternative channels such as the backward linkages between the extractive industry and the rest of the economy, or the consequences at local level where many of the negative effects, such as pollution or population displacement, occur.

In this paper, we fill this gap in the literature by investigating whether natural resources benefit local communities and exploring the economic mechanisms for this to happen. We use the case of Yanacocha, a large gold mine in Peru, as a testing ground and evaluate the effect of the expansion of the mine on the income and welfare of the local population. In contrast to the existing literature, we focus on the general equilibrium effects of the mine's demand of local inputs as the main transmission channel.

In order to inform the way to look at the data, we first develop a spatial general equilibrium model based on Fujita and Krugman (1995). In this setup, there is a single city surrounded by a rural hinterland, specialized in the production of manufactured goods and food, respectively. Both areas engage in costly intra-regional trade, and are also able to import goods from other regions. The model treats the expansion of the mine as a demand shock on local labor in the city and delivers two testable predictions. First, if the mine expansion increases city's real income, then the agricultural real income also increases. The effect wanes the longest the distance to the mine's input market due to transportation costs. Second, there is an increase on the relative price of locally produced food due to the relative larger demand. The increase on agricultural prices transmit the income gains to the rural population, not directly supplying goods or services to the mine.

We test these predictions using household data from repeated cross sections. The data comes from household surveys - representative of the region where the mine is located - for the period 1997 to 2006. We exploit two sources of variation to identify the effects of the mine expansion. First, we use the significant increase on Yanacocha's local purchases and wage bill since 2000. This phenomenon was driven by the increment on gold extraction and the implementation of a corporate policy directed at increasing local procurement and employment. Second, we use the distance from the household location to Cajamarca city - the mine's local supplying market - as a source of heterogeneous exposure to the mine's expansion.

We find a positive effect of the mine's activities on real income in the city and surrounding areas. The effect decreases monotonically with distance and becomes insignificant beyond 100 km from the city. Our estimates suggest that a 10 percent increase in the mine's local purchases and wage bill is associated with an increase in real income of 1.7. We also observe increments on the relative price of local food crops, such as potatoes and maize. This is consistent with the effect on real income being driven by a market mechanism and explains how the rural population benefitted from the demand shock in the city markets.

Nonetheless, in the presence of locational externalities associated to the mine (such as pollution or crime), increments on real income do not imply an increase on welfare. Similarly, the average increase on income may hide negative re-distributional effects, for example if poor households are unable to benefit from the mine expansion. To explore these questions, we first follow Roback (1982) who suggests using house prices as a measure of welfare.[1] We find evidence of increments on house rental prices, in line with the observed increase on real income, which we interpret as evidence of net welfare gains due to the mine expansion. Second, we test for changes in self-reported measures of health and crime, but we do not find evidence that they have worsened with the expansion of the mine. Thirdly, we analyze the re-distributional impacts of the mine expansion. Using quantile regressions,

[1] For an example of an empirical application, see Greenstone and Moretti (2003).

we find that the average household income at the bottom of the income distribution has experienced an increase of real income similar to that observed for richer households. We also find evidence of poverty reduction associated to the mine expansion.

Finally, we investigate whether the observed phenomena is driven by the revenue windfall to local governments or by the mine's backward linkages. This is a relevant question because of the importance attached to the increment on public expenditure as one of the main benefits from natural resource exploitation. We show that local public spending increases due to the expansion of the mine. However, we find no evidence that this expansion of the public sector contributes to the observed increase in real income or welfare.[2] Instead, the observed phenomena seems to be driven entirely by the expansion of the mine's local purchases and employment. To the best of our knowledge, there is not previous empirical work contrasting the relative importance of these two mechanisms.

In sum, our empirical results suggest that the expansion of the mine increased real income and welfare of the local population. The effect seems to be driven by the demand shock and its multiplier effect, associated to the mine's backward linkages not to the fiscal revenue windfall. The gains are transmitted to residents in the rural area, not directly selling inputs to the mine, due to the existence of trade within the economic region. The increase on the relative price of locally produced food crops is evidence of this transmission channel.

A main policy implication of our findings is that, even in the presence of weak governments, natural resources can benefit local populations if backward linkages between the extractive industry and the rest of the economy are strong enough. However, as suggested by the model, this recommendation hinges on the existence of economic integration in the regional economy (labor mobility and trade) as well as the pre-existence of good and labor markets able to supply local inputs to the extractive industry.

The lessons to be drawn from this paper are not exclusive to extractive industries, they could apply to any business venture that creates a strong demand shock in a relatively poor area. There are, however, at least two features that make large-scale mining a different and interesting case to study.

[2]Using the case of Brazil, Caselli and Michaels (2009) also find evidence suggesting that the oil revenue windfall to municipalities did not translate to increases on local income.

First, mining is a relevant industry for developing countries.[3] According to the World Bank, around 60 countries in the world have a mining sector large enough, in terms of GDP or exports, to be a main driver of economic growth and poverty reduction (World Bank, 2002). However, understanding how to transform mining wealth into better living standards remains an unsolved policy question. Second, there are well-documented negative local effects associated to mining. This increases the need to assess both direct and indirect benefits on local communities to inform industrial and development policies.

This paper contributes to the literature studying the effect of natural resources on economic development and to the literature examining how regional markets respond to demand shocks.[4] Three features distinguish our paper from previous work. First, we use a microeconomic approach where the household is the unit of observation, as opposed to counties or districts. The comprehensiveness of the data allows us to explore in more detail the mechanisms generating the spillover effects and to assess the net impact on resident's welfare. Second, we show the importance of backward linkages as a channel to improve local conditions. The mechanism we uncover hints at an alternative explanation for the failure of natural resources to promote development: few linkages with local markets. This explanation relates to a technological feature of extractive industries and contrasts to the political economy argument used to explain the natural resource curse. Finally, we

[3]Developing countries play an important role in global mining. They are the largest producers of key minerals like iron, copper and aluminium, and have almost three quarters of the world mineral resources. They account for most of the recent growth on mining production and attract an even increasing proportion of exploration spending (Humphreys, 2009).

[4]There is not conclusive evidence of the effect of natural resources on development. Earlier studies use cross country data and find negative effects (Sachs and Warner, 1995; Sachs and Warner, 2001). However other papers using alternative measures of resource abundance fail to find a negative relation (Lederman and Maloney, 2003; Sala-i-Martin et al., 2004). Using a microeconomic approach, Michaels (2006) finds that oil abundance in U.S. counties caused specialization, but that these costs were offset by population growth and better education. In balance, oil-rich counties enjoyed higher income per capita without increase in inequality. Black et al. (2005) explore the effect of a coal boom in U.S. counties and find modest positive spillover effects on employment. More recently, Caselli and Michaels (2009) find no evidence of positive effect of oil extraction on income in Brazilian municipalities. In contrast, the literature on the impact of demand shocks (such as large industrial plants, construction works, casinos or military bases) find some evidence of positive regional effects on wage bill and employment (Carrington, 1996; Hooker and Knetter, 2001; Evans and Topoleski, 2002), and resident's welfare (Greenstone and Moretti, 2003).

embed the study of natural resources (and local demand shocks) in the analytical framework of a new economic geography model. This allows us to explicitly explore general equilibrium effects and to introduce transportation costs, agglomeration economies and migration which are relevant factors when studying a regional economy.

The remainder of the paper is organized as follows. Section 2 develops the analytical framework we use to explore the effect of the mine on a regional economy. Section 3 presents background information about Yanacocha and the relevant geographic area. Section 4 describes the data and identification strategy. Section 5 presents the main empirical results related to the model predictions. Section 6 explores alternative explanations of the observed phenomena. Section 7 presents robustness checks and section 8 concludes.

2 Analytical framework

In this section we present a stylized model to shed light on the general equilibrium effects of the expansion of a mine on a regional economy. We use the mono-centric city model developed by Fujita and Krugman (1995) and extend it by including an export sector and interregional trade. In this framework, the mine is a foreign-owned net exporter that uses local labor as a production factor but sells all its output in international markets.[5]

We treat the expansion of the mine as a shock in the demand for local labor and explore the effect on factor and good prices, and ultimately on real income. The model allows us to analyze the spatial distribution of the effects as well as the interaction of different economic sectors.

The model stresses the role of backward linkages, rather than alternative channels such as forward linkages, agglomeration economies, public spending or technology spillovers.

2.1 Environment

We consider a mono-centric spatial structure à la von Thünen, as depicted in Figure 1. In this spatial configuration, the regional economy is a long narrow line with a single city surrounded by a rural hinterland that extends

[5]This setup resembles the characteristics of the mining activities of Yanacocha and other large modern mines.

from $-f$ to f, where f represents the endogenously determined agricultural frontier. The land is of homogenous quality with one unit of land per unit of distance. Each location along the line is denoted by r. For simplicity, we drop it when referring to the city ($r = 0$).

The economy has three productive sectors: manufacturing, agriculture and an export sector. The manufacturing sector produces a large number of varieties with an increasing returns to scale technology and has a monopolistic competition structure. In contrast, the agricultural sector is perfectly competitive and produces a single homogeneous good, namely food. The export sector produces two goods: a manufactured commodity and the natural resource, which are not consumed locally. Consumers in the region can import consumption goods from other regions. Import and export goods are homogenous and their prices are set in external markets.

All manufacturing, production of the export goods and import activity takes place in the city, while the rural hinterland specializes in food production. We take this spatial system as given.[6] There are (potentially) different factor and good prices in each location r, since we assume trade within the region is costly. We use an iceberg transportation cost such that if 1 unit is transported over a distance d, only $e^{-\tau d}$ units arrive. The parameter τ represents the transportation cost and is the same for all types of goods.[7]

The region has a given population of N workers who supply inelastically one unit of labor.[8] Intra-regional migration is costless and workers can freely move among locations. They live where they work, so we rule out the possibility of commuting. The free mobility assumption implies that the population allocation within the region is endogenous. In particular, the size of the city L will depend on the spatial distribution of real wages.

Consumers All consumers share the same homothetic preferences $U = X^{\mu} A^{\alpha} M^{1-\alpha-\mu}$, where X is a composite index of the domestically produced

[6] Fujita and Krugman (1995) show that this configuration can be an equilibrium outcome, if some conditions are met. In particular, if the population is relatively small, if manufacturing goods represent a large enough proportion of consumers' expenditure or if consumers have a sufficiently strong preference for varieties.

[7] The model predictions are similar if we assume different transportation costs for agricultural, manufactured and import goods.

[8] This assumption is very important since it precludes the possibility of inter-regional migration. However, the assumption is not a very strong one, if we take into account evidence in the next section that shows very low levels of inter-regional migration flows in Cajamarca over the analyzed period.

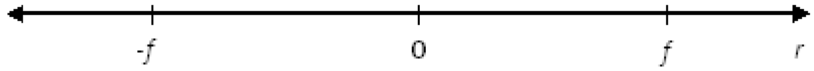

Figure 1: The monocentric economy

manufactured goods, A is the agricultural good or food and M is the import good. The quantity index X is a CES function of the consumption of each available variety:

$$X = \left(\int_0^v c_i^\rho di\right)^{\frac{1}{\rho}},\ 0 < \rho < 1,$$

where c_i is the consumption of each individual variety and v is the range of varieties. The parameter ρ denoted the taste for variety and $\sigma = \frac{1}{1-\rho}$, the elasticity of substitution between varieties.

The representative consumer's budget constraint in location r is given by

$$Y(r) = p_a(r)A(r) + p_m(r)M(r) + \int_0^v p_i(r)c_i(r)di,$$

where Y is the consumer's income, p_a is the price of food, p_m is the price of the import good and p_i is the price of variety i. Note that, due to transport costs, prices and quantities are location-specific.

Consumers maximize their utility given the budget constraint. We can solve the problem in two steps.[9] First we minimize the cost of attaining X and obtain the compensated demands for each variety $c_i(r) = \left(\frac{p_i(r)}{G(r)}\right)^{-\sigma} X$ where $G(r) = \left[\int_0^v p_i(r)^{1-\sigma}di\right]^{\frac{1}{1-\sigma}}$ is the price index for the manufacturing composite. Then, we allocate the available budget between A, M and X using the utility function parameters α and μ. The demand for the agricultural and import goods are $\alpha\frac{Y(r)}{p_a(r)}$ and $(1-\alpha-\mu)\frac{Y(r)}{p_m(r)}$ respectively, while the demand of the composite X is $\mu\frac{Y(r)}{G(r)}$. Combining this expression with the compensated demand we obtain the uncompensated demand for each variety:

$$c_i(r) = \mu Y(r)\frac{p_i(r)^{-\sigma}}{G(r)^{-(\sigma-1)}}. \quad (1)$$

[9]For a detailed exposition see Fujita et al. (1999, pp. 46-48).

Furthermore, from the indirect utility function we can obtain the cost of living in a specific location r:

$$P(r) = G(r)^{\mu} p_a(r)^{\alpha} p_m(r)^{1-\alpha-\mu}. \tag{2}$$

Manufacture producers The local manufacturing industry uses an increasing returns to scale technology with labor as the only production factor. In particular a firm requires $l_i = F + \beta x_i$ units of labor to produce x_i units of a variety of the manufacturing good. The presence of increasing returns to scale implies that each variety is produced by a single firm. There is free entry and exit to the industry, hence in equilibrium manufacturing firms make zero profits.

A firm's profit is equal to $p_i x_i - w l_i$, where $x_i = c_i$ as in equation (1) and w is the wage in the city.[10] The resulting profit-maximization price is

$$p_i = \frac{\sigma}{\sigma - 1} \beta w. \tag{3}$$

Note that this result implies $p_i = p$, i.e. all varieties have the same price, since they all share the same marginal cost β. It follows that the manufacturing price index becomes $G = p v^{\frac{1}{1-\sigma}}$ and the location-specific individual demand for each variety (1) as

$$c(r) = \frac{\mu Y(r)}{p(r) v}. \tag{4}$$

From the zero-profit condition we obtain a firm's individual supply $x = \dfrac{F(\sigma - 1)}{\beta}$ and its labor demand $l = F\sigma$.

Agricultural production and rural hinterland Farming uses a fixed proportion technology. It requires 1 unit of land and c_a units of labor to produce 1 unit of food. Food is consumed locally and the surplus, a proportion $1 - \alpha$, is sold to the city.[11]Landlords live in their landholdings, hence all the agricultural income is consumed locally. For a given frontier f the total demand of labor from the agricultural sector is $2c_a f$ and total food supply to the city is $2(1-\alpha) \int_0^f e^{-\tau r} d$

[10]In the city, given the manufacturing sector zero-profit condition, the wage w is also the average income.

[11]This follows from the representative consumer's maximizing food consumption.

Food is traded in the city at a price p_a. However due to transportation costs, the price of food received by farmers decreases with distance to the city, i.e. a farmer in location r receives $p_a(r) = p_a \ e^{-\tau|r|}$ where $|r|$ measures the distance to the city. This means that the opportunity cost of consuming food is lower the greater the distance to the city.

In contrast, since the manufactured goods and import need to be shipped from the city to rural areas for their consumption, the corresponding prices increase with distance. In particular, in location r the price of the manufacture composite is $G(r) = Ge^{\tau|r|}$ and the price of imports, $p_m(r) = p_m e^{\tau|r|}$. Taken together, these results imply that the price index in location r is

$$P(r) = Pe^{(1-2\alpha)\tau|r|}. \tag{5}$$

Under the assumption that locally produced food represents less than half of a household's expenditure ($\alpha < 0.5$), the cost of living increases with distance.[12]

Since landlords use c_a units of labor to produce 1 unit of food, their rents are $R(r) = p_a \ e^{-\tau|r|} - c_a w(r)$, where $w(r)$ is the wage rate in location r. The hinterland extends around the city up to the point where rents are zero, or $w(f) = \dfrac{p_a e^{-\tau f}}{c_a}$. Costless migration implies that in equilibrium the real wage is the same in all locations. In particular we compare the real wage in the city and in the agricultural frontier to obtain the following wage equalization condition[13]:

$$w = \frac{p_a e^{-2(1-\alpha)\tau f}}{c_a}. \tag{6}$$

Finally, we derive total demand for manufactures and import goods from the rural hinterland. Note that a proportion μ of the local income is consumed in the manufactured good while the remaining $1 - \alpha - \mu$ is spent on imports. In the rural hinterland, the total income -distributed between landlord and workers- corresponds to the price of the agricultural output $p_{a(r)}$. Since landlords and agricultural workers share the same homothetic

[12]We require farmers to trade a significant part of their production. It does not mean that the consumption basket of rural farmers is not mainly food, since many of the imported goods can be foodstuff produced elsewhere. Results do not hinge on this assumption.

[13]Real wage equalization and the price index defined in (5), give us a nominal wage in location r equal to $w(r) = we^{(1-2\alpha)\tau|r|}$.

preferences, we can treat their demand as the one of a representative consumer. Replacing the values of local prices and income, we obtain the total demand of each manufactured good variety and import goods from the rural hinterland:

$$\text{Manufactured goods:} \quad 2\mu \frac{p_a}{pv} \int_0^f e^{-2\tau r} dr \tag{7}$$

$$\text{Imported goods:} \quad 2(1-\alpha-\mu) \frac{p_a}{p_m} \int_0^f e^{-2\tau r} dr. \tag{8}$$

Export sector There are two types of firms in the export sector: producers of a tradeable commodity and a mine that extracts the natural resource. Both use labor as their input and sell their output in perfectly competitive external markets.

In contrast to the manufacturers supplying the domestic market, the commodity producers face decreasing returns to scale. In particular, their production function is $E(L_e) = L_e^{\varepsilon}$, where E is the amount produced of the commodity, L_e is the size of the workforce employed in the industry and $\varepsilon < 1$. The industry profit is $p_e E(L_e) - wL_e$, where p_e is the price of the commodity. Solving for the profit maximization problem we obtain the unconditional labor demand function:

$$\text{Labor demand: } L_e(w) = \left(\frac{\varepsilon p_e}{w}\right)^{\frac{1}{1-\varepsilon}}. \tag{9}$$

Note that $\varepsilon p_e E$ is paid, as wages, to local workers; while the industry profits $\pi_e = (1-\varepsilon) p_e E$ are distributed among the firms' shareholders. For simplicity we assume that all shareholders reside in the city and have similar preferences than workers.

The mine uses a fixed proportion of local labor in its production process. The demand of local labor is θS, where S is the mine's production and $\theta < 1$ is exogenously determined. Note that θ captures the extent of backward linkages between the mine and the local economy. The mine is foreign-owned and there are no taxes, hence all the profits are remitted abroad.

2.2 Equilibrium

The instantaneous equilibrium is given by the solution of the following system of non-linear equations:

$$\frac{F(\sigma-1)}{\beta} = \frac{\mu}{pv}\left(Lw + 2p_a\int_0^f e^{-2\tau r}dr + \pi_e\right) \quad (10)$$

$$2(1-\alpha)\int_0^f e^{-\tau r}dr = \frac{\alpha}{p_a}(Lw + \pi_e) \quad (11)$$

$$L = vF\sigma + L_e(w) + \theta S \quad (12)$$

$$w = \frac{p_a e^{-2(1-\alpha)\tau f}}{c_a} \quad (13)$$

$$p_e E(w) + w\theta S = (1-\alpha-\mu)\left(Lw + 2p_a\int_0^f e^{-2\tau r}dr + \pi_e\right) \quad (14)$$

$$p = \beta w\frac{\sigma}{\sigma-1} \quad (15)$$

$$N = L + 2c_a f. \quad (16)$$

Equation (10) represents the market equilibrium of each variety of the local manufacture. The total demand on the right hand side is proportional to the regional income which is composed by the wages paid to city workers, agricultural income and the profits from the commodity export firms. Equation (11) is the food market equilibrium condition while equation (12) represents the equilibrium in the city's labor market. Note that the labor demand in the city comes from local manufacturers, commodity export firms and the mine.

Equation (13) is the wage equalization condition (6) obtained from the assumption of costless migration. Equation (14) represents the current account's balance: in equilibrium the value of exports minus the net factor payments should be the same as the value of total imports.[14] Equation (15) is the manufacturers' profit maximization condition (3). Finally, equation (16) represents the population constraint, where the size of the regional population equals the number of workers in the city and in the agricultural sector.

We solve the system analytically. First note that, using (12) and (16), we can write the number of manufacturing firms, v, as:

$$v(w,f) = \frac{N - 2c_a f - L_e(w) - \theta S}{F\sigma}. \quad (17)$$

[14]We assume a fixed exchange rate equal to 1.

Using this result, replacing (14) into (10) and re-arranging, we obtain:

$$N - 2c_a f - \frac{(1-\alpha-\mu)\,\varepsilon + \mu}{(1-\alpha-\mu)\,\varepsilon} L_e(w) - \frac{1-\alpha}{1-\alpha-\mu}\theta S = 0. \qquad \text{(AA)}$$

Note that expression (AA) defines an upward sloping curve in the space (f, w). This curve has a positive intercept and also an upper bound $\overline{f}$ on the values that f can adopt.[15]

Second, using (9), (13) and (16), solving for the integral over f and rearranging, we can rewrite the equilibrium condition in the food market (11) as:

$$2c_a[\frac{(1-\alpha)(1-e^{-\tau f})}{\alpha e^{-2(1-\alpha)\tau f}} + f] - N - \frac{(1-\varepsilon)}{\varepsilon} L_e(w) = 0. \qquad \text{(BB)}$$

This expression defines a downward sloping curve in the space (f, w), with a lower bound $\underline{f}$ on the values of f.[16]

Intuitively, the curve AA is upward sloping because with a larger rural hinterland (higher f), the city has a smaller population and hence the labor supply is smaller. This translates into higher wages in the city. In contrast, curve BB is downward sloping because higher wages in the city increase agricultural wages (due to costless intra-regional migration). In turn, this means that the marginal farmer makes losses and hence the agricultural frontier shrinks (smaller f).

To solve the model we plot curves AA and BB (see Figure 2). The intersection of both curves at point Q provides the unique pair (f^*, w^*) that solves the system and defines the equilibrium.[17]

2.3 The Effects of the Mine's Expansion

We are now ready to use the model to perform a thought experiment and analyze the effect of the mine's expansion on the regional economy.

We simulate the expansion of the mine as an increase on S, the amount of

[15] To see this, note that $\overline{f} \equiv \lim_{w\to\infty} f = \frac{1}{2c_a}\left(N - \frac{1-\alpha}{1-\alpha-\mu}\theta S\right)$ and that $\lim_{f\to 0} w = L_e^{-1}\left(2c_a \overline{f} \frac{(1-\alpha-\mu)\varepsilon+\mu}{(1-\alpha-\mu)\varepsilon}\right)$, where L_e^{-1} is the inverse function of $L_e(w)$.

[16] Note that $\underline{f} \equiv \lim_{w\to\infty} f = g^{-1}(N)$, where $g^{-1}(\cdot)$ is the inverse function of $g(f) \equiv 2c_a[\frac{(1-\alpha)(1-e^{-\tau f})}{\alpha e^{-2(1-\alpha)\tau f}} + f]$.

[17] See Appendix A.1 for a formal proof.

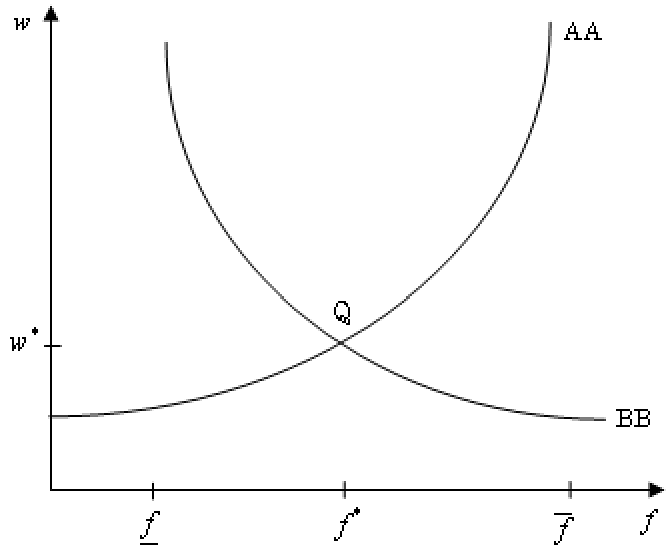

Figure 2: Model equilibrium

natural resources extracted by the mine. The immediate effect of this shock is to increase demand of labor and shift the curve AA upwards as depicted in Figure 3. To understand this shift note that f defines the distribution of people between city and rural countryside, and thus the labor supply in each area. Hence, keeping the labor supply fixed, increments on the demand for labor in the city due to the mine expansion would increase local wages.

As we see in Figure 3, the equilibrium moves from point Q to $Q\prime$ with a higher wage in the city but smaller f. The reduction on f implies a smaller extension of agricultural frontier and a larger city population. Since N is assumed to be fixed, the grow of the city requires migration of agricultural workers to the city.[18] In this model, the increase on wages - triggered by the mine expansion - is partially offset by migration to the city.

Crowding out The expansion of the mine, and the subsequent wage increment, displaces other economic activities. First, the reduction on the agricultural frontier implies a smaller agricultural production. Second, the higher wage reduces the employment, output and profits of the export sec-

[18]The assumption of a fixed N is important to obtain the model predictions. If we assume instead costless inter-regional migration, the effect of the mine expansion on real income would be zero.

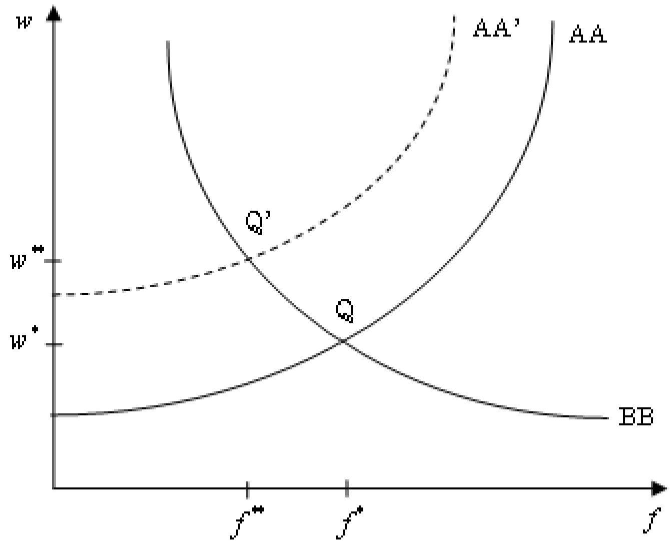

Figure 3: The expansion of the mine

tor which uses labor as its only production factor. This result is similar in flavor to the "Dutch disease", commonly found in macroeconomic models of resource boom (Corden and Neary, 1982). In this case, the crowding out is driven by increase on the relative price of a production factor.

Nonetheless, the effect on the local manufacturing industry could be either positive or negative. On the one hand, the growth of the city population favors the increase on the number of firms, v, due to the increasing returns to scale. On the other hand, this positive effect is offset by the expansion of the mine's labor demand, which reduces the labor available for manufacturing firms. The net effect depends on the extent of the mine's backward linkages, θ. In particular, if it is sufficiently high, v would increase, because the multiplier effect would make the city population grow enough to satisfy the additional labor demand from the mine without harming the manufacturing industry.[19]

[19]See Appendix A.2 for a formal derivation of the condition.

Effect on real wages From (2), we can obtain an expression for the real wage in the city:

$$\omega = \frac{w}{P} = \left(\frac{w}{G}\right)^{\mu}\left(\frac{w}{p_a}\right)^{\alpha}\left(\frac{w}{p_m}\right)^{1-\alpha-\mu}, \tag{18}$$

where $G = pv^{\frac{1}{1-\sigma}}$ and $\sigma > 1$.

Note that the wage relative to the price of food and import goods ($\frac{w}{p_a}$ and $\frac{w}{p_m}$) increases as a consequence of the expansion of the mine.[20] These results highlights two sources for real wages to increase: the increase on the payments to labor relative to land rents, due to the additional demand; and the access to relatively cheaper import goods. However, the effect on real wages is ambiguous, because these gains may be offset by higher price of manufactured goods (G), if the number of firms reduces.[21]

2.4 Testable predictions

The previous results provide the basis for the two testable predictions of the model regarding the effect of the mine expansion on real income and relative prices.

2.4.1 Effect on real income

The income in the city and rural hinterland comes from different sources. In the city, the income is composed by wages and the dividends from export firms. In contrast, in the rural hinterland it is proportional to the value of the agricultural output $p_a(r)$. Taking that into account, we can write the average real income in location r as:

$$y(r) = \begin{cases} \omega + \frac{\pi_e}{PL} & \text{if } r = 0 \\ \frac{1}{k}\frac{p_a}{P}\ e^{-2(1-\alpha)\tau|r|} & \text{if } r \neq 0, \end{cases} \tag{19}$$

[20] To see this note, from (13), that $\frac{w}{p_a}$ is decreasing on f and hence increases with S. Additionally, given that the price of imports p_m is exogenously determined and w increases, $\frac{w}{p_m}$ also increases.

[21] We can, however, identify two plausible cases in which the real wage increases: when v increases, and when v decreases but local manufactures represent a small proportion of the budget (small μ). See Appendix A.3 for a formal derivation of these conditions.

where $y(r)$ is the average real income, P is the general price index in the city, and k is the unknown, but constant, population size in each rural locality.[22]

The effect of the mine expansion on the income of city residents is, a priori, unclear and remains an empirical question. On the one hand, it increases due to higher wages. On the other hand, it decreases due to reduction on the profit of export firms, crowded out by the mining sector, and the (potential) increase on the price of the manufactured good due to the smaller number of firms.

We focus on the case in which the real income in the city increases with the mine expansion. In this case, the model provides clear testable predictions regarding the relation between income in the city and rural hinterland and relative prices.

Proposition 1 *If the real income of city residents increases, then the real income of rural inhabitants must also increase.*

Intuitively, the growth of urban income combined with the larger city population increases the demand for food. This demand shift, combined with the reduction on food supply due to the smaller agricultural frontier, necessarily results on an increase in the price of food and on higher income in rural locations.

To explore the effects on welfare, note that the indirect utility of a household in location r is $V(r) = \mu^{\mu}\alpha^{\alpha}(1-\mu-\alpha)^{1-\mu-\alpha}y(r)$. Hence, the real income becomes a sufficient statistic for welfare, however this result does not hold in presence of externalities from the mine. For that reason, in the empirical section we explore different measures of welfare in addition to real income.

In this model, the shock of demand created by the mine expansion directly affects the city's labor and goods markets. The effect is transmitted to the surrounding rural areas due to the trade links between the city and its rural hinterland. Since intra-regional trade is limited by the existence of transportation costs, a relevant question is how the effects of the mine expansion are distributed spatially in the economic region.

Proposition 2 *When the mine expansion increases real income, the effect decreases monotonically with distance to the city.*

[22] Recall that in each farm there are c_a workers and a landlord, but it does not correspond to the total population since a landlord can also be a worker.

The model predicts heterogeneous effects by distance, because transportation costs reduce the land rents for farmers in farther locations and absorb part of the gains on real income. In the extreme, beyond the agricultural frontier, the effect is nil since there is no trade with the city. In the empirical section, we use this insight as a key element of the identification strategy. In particular, we obtain the distance from a locality to the city, and use it as a measure of different exposure of households to the mine expansion.

2.4.2 Effect on relative prices

The increase of real income, and its transmission to agents not directly employed by the mine, are driven by changes in the relative prices of labor and food. To explore this mechanism further, we use the model to obtain predictions regarding the effect of the mine expansion on relative prices within the economic region. We focus on two types of goods that are possible to identify in the data: the agricultural good and the import good.

Proposition 3 *If the real income of city residents increases due to the mine expansion, then* $\frac{p_a(r)}{P(r)}$ *also increases. The effect on* $\frac{p_m(r)}{P(r)}$ *is ambiguous.*

The relative price of food in the region increases because it is proportional to the real agricultural income. In contrast, the effect on the relative price of the import good is ambiguous since G, one of the components of the price index may increase or decrease. However, if local manufactures represent a small proportion of the consumption basket, then the relative price of imports would actually decrease due to the increase on price of food. A corollary of this result is:

Proposition 4 *if the real income of city residents increases due to the mine expansion, then the increase on* $y(r)$ *is greater than the increase on* $\frac{Y(r)}{p_a(r)}$*, where* $Y(r)$ *is the nominal income in location* r*.*

Note that $\frac{Y(r)}{p_a(r)} = \frac{y(r)}{p_a(r)/P(r)}$. It is straightforward to see that the change on $\frac{Y(r)}{p_a(r)}$ is smaller than the change on $y(r)$ since $\frac{p_a(r)}{P(r)}$ also increases due to the mine expansion.

This result sheds light on the transmission channel of the effect from the city to the rural hinterland. Since rural income increments are driven for higher food prices, the income relative to the price of food should not vary as consequence of the mine expansion. In the city, the increment may be positive, but still smaller than the increase on y.

3 The Case of Yanacocha Gold Mine

Peru has a long tradition as a mining country and ranks among the top producers of minerals in the world.[23] In the late 1990s it experienced a mining boom and the sector expanded significantly. In the relevant period for our analysis, from 1997 to 2006, the mining product grew at an average annual rate of 7.4% while the economy growth rate was around 3.5%. As a consequence, the share of the mining sector in GDP went from 3.8% to 5.2% (Banco Central de Reserva del Perú, 2009). The mining growth has been mostly driven by the opening and expansion of large mining operations in gold, copper and silver.[24]

One of the largest new mines is Yanacocha. The mine is privately owned and none of the investors resides in the locality.[25] Yanacocha is the second largest gold mine in the world, representing around 45% of the total Peruvian gold production, and the most profitable. The mine extracts gold from open pits and uses a very capital intensive technology. All production is exported as gold bars, with no other processing or added value. This feature precludes the creation of forward linkages.

The mine is located in the department of Cajamarca in the North Highlands of Peru. The department of Cajamarca is mostly rural and relatively poor. In 2006, Cajamarca was the eighth department (among 24) in terms of incidence of poverty, with 63.8% of the population living below the poverty line (INEI, 2007). The mine is located less than 50 km. away from Cajamarca city, the department's capital and largest urban settlement (see Figure 4 for a map of the spatial configuration). Due to the proximity to the city, most workers and local suppliers live there. This feature facilitates

[23]In 2006 it was the first producer of gold, zinc, silver, lead and tin in Latin America and among the top five producers of those minerals in the world (Ministerio de Energía y Minas, 2006).

[24]Some of the new mines are Yanacocha (gold), Pierina (gold) and Antamina (silver and copper).

[25]The mine shareholders are: Newmont Mining Co., with more than 50% of the shares, Compañia de Minas Buenaventura (more than 40% of the shares) and the World Bank (around 5% of the shares). Newmont Mining Co. is a U.S. based company, while Compañia de Minas Buenaventura has its headquarters in Lima, Peru's capital city.

our analysis, because the city becomes the geographical market where the mine purchases local goods and services.[26] There are other three major cities in the surroundings (Trujillo, Chiclayo and Chachapoyas) and some small and medium mines in the south of Cajamarca.[27]

3.1 Mine Production and Local Procurement

Yanacocha started operations in 1993, and its production increased steadily over time. In 2001, it expanded its production due to the opening of new gold deposits Figure 5 shows the expansion of the mine production since 1997 and the change of trend since 2001 associated to the expansion of the production capacity.

The growth on the mine size has required a steady increase on the purchase of production inputs such as explosives, machinery, oil, laboratory services, limestone, etc. As we will see below, the mine expansion coupled with a shift in procurement policies has increased Yanacocha's local purchases and employment.

Yanacocha procures its inputs in three ways: purchasing goods to suppliers, directly hiring workers and contracting out services. In terms of origin, Yanacocha classifies the supplying firms as local and non-local depending of where they are registered for tax purposes and the origin of their shareholders.[28] In particular, for a firm to be considered local it has to pay taxes in Cajamarca and have at least 50% of Cajamarquinos among their shareholders. For individuals, the requirement is to be born in Cajamarca or to demonstrate continuous residence for at least five years in Cajamarca.[29]

In the first years of operation, Yanacocha had little economic interaction with the city of Cajamarca and its surroundings, with only a very small proportion of inputs bought locally (Kuramoto, 1999). For example, in 1998 the value of purchases from Cajamarca represented 3% of total good purchases and 15.5% of services. In contrast, purchases from Lima were 77.5% and 72.6% of goods and services respectively. The rest was bought

[26]Yanacocha classifies its purchases by origin as local (sourced in Cajamarca), national (bought from the rest of the country, mostly Lima) and foreign.

[27]These mines use more traditional technologies, have smaller scale and are more labor intensive than Yanacocha.

[28]The non-local category includes Lima, rest of Peru, and other countries.

[29]This information was obtained from internal memorandums and interviews with Yanacocha representatives.

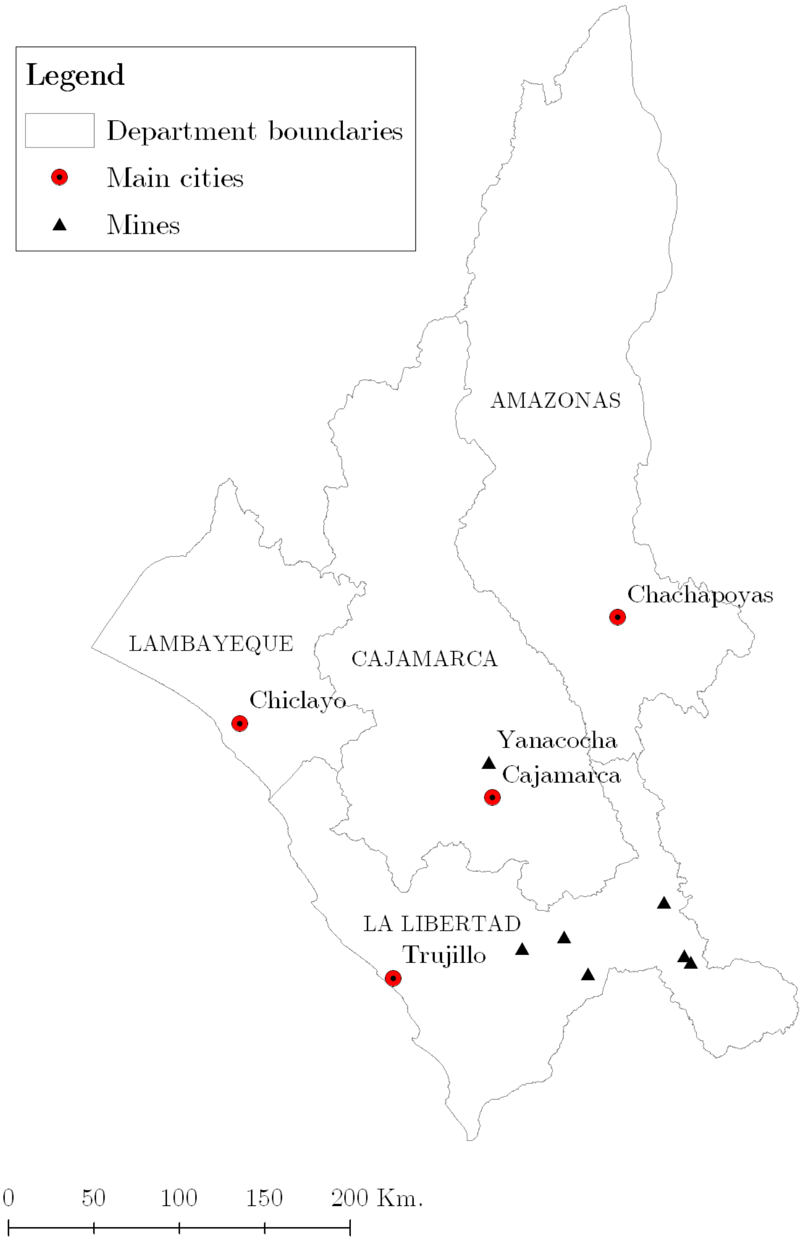

Figure 4: Location of mine and main cities

abroad or in the rest of the country. There were also very few local firms among the top mine suppliers.[30]

A shift in procurement policy Since late 1990s, the mine started a shift in its procurement and employment policies aimed to increase the participation of local firms and workers. The policy change was promoted by the International Finance Corporation, one of Yanacocha's shareholders and Newmont's main lender, as a way to increase the economic impact of the mine in the region and minimize the risk of conflict with the local population (Jenkins et al., n.d.).[31]

The most important feature of the mine procurement policy was to give priority to local suppliers and workers in competitive bids and recruitment.[32] For example, a 2006 policy memorandum by the mine aiming at outsourcing services such as cleaning, waste management or transport, establishes that "in competitive bids with non-local companies, the Company has to prioritize the hiring of Cajamarquino small and medium enterprizes". Similarly, the memorandum specifies that the mine should ensure that at least 60% of personnel of the mine contractors reside in Cajamarca. This shift in policy becomes apparent in the evolution of local purchases and wage bill (including bonuses) that show a significant increase since 2000 (see Figure 5).

What does the mine buy locally? The local purchases consist of relatively simple services and goods, with low quality requirements and not considered vital to the mine operations. The expenditure on services is larger that the expenditure on goods.[33] The main local services contracted locally include demolition, haulage and transport of material to leach pads, and cleaning and reforestation activities. More specialized services such as transport of dangerous materials, consulting, engineering, laboratory analysis, etc., are supplied by firms located in Lima or abroad. In the case

[30] Among the top 20 suppliers, there was only one local firm. In the case of service providers, there were four local firms among the top contractors.

[31] IFC is promoting policies to increase supply chain linkages in other mining, oil and gas projects. For example, a similar approach is being development in a new mine operated by Newmont in Ghana.

[32] Yanacocha has also promoted quality certification among its suppliers, but it has not engaged in direct transfer of technology or training.

[33] For example, in 2004 the expenditure on services was around nine times greater than the expenditure on goods.

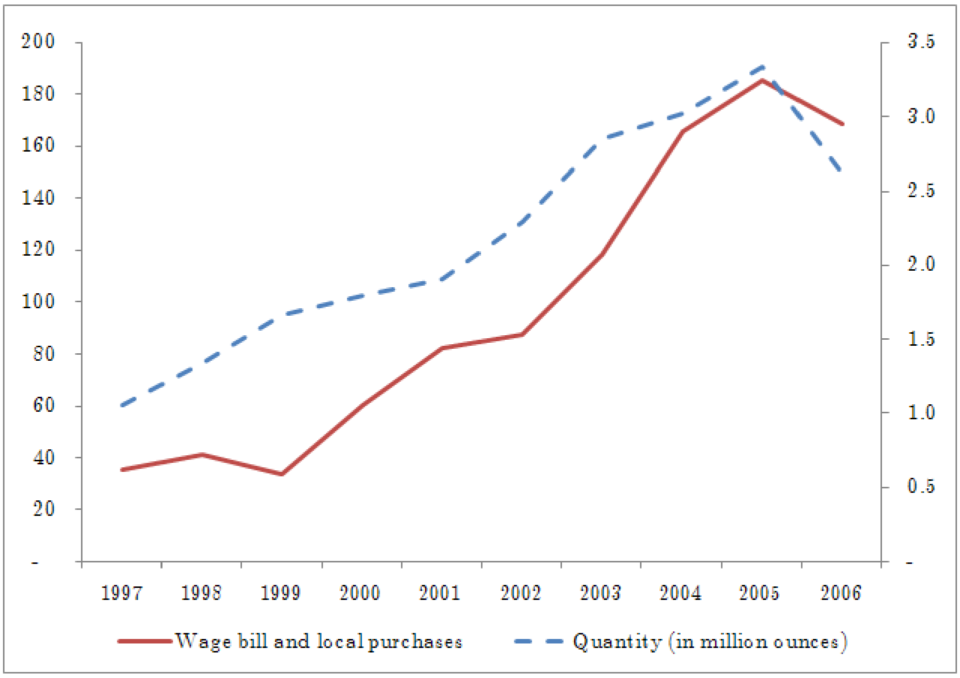

Figure 5: Evolution of Yanacocha's local procurement and production

of goods, local purchases consists mostly of construction materials, basic hardware, cleaning materials, vehicles spare parts chemical products, office furniture, etc.

Even though local purchases and wages would imply different channels through which the activity of the mine affects household incomes (e.g. direct effect of wages or development of backward linkages), the distinction is not as clear-cut as it seems. The mine hires some workers directly; however, most workers located in the mine site are employed by contractors who provide services to the mine.

Figure 6 shows the evolution of the number of workers in the mine, directly employed by Yanacocha or indirectly through contractors. In the period 1997-2006 the number of total workers increases from 4,000 to 14,000, a substantial proportion of the active population of Cajamarca. To put this numbers in context, the mine's workforce represented around 12% of the active population of the city of Cajamarca in 2001 and 20% in 2005. Consistent with the mine policies, 60% of the workers are locals and almost all of them live in Cajamarca (Minera Yanacocha, 2006)

The increase of mine employment occurred in 2000 and was driven mainly by workers indirectly hired through contractors. Contractor workers represent a significant proportion of the mine's workforce. In average they represent 78.5% of the total mine workers. Workers that are hired directly by Yanacocha, tend to be skilled (engineers, accountants, technicians, secretaries, assistants, supervisors) as opposed to low-skilled workers employed through contractors.

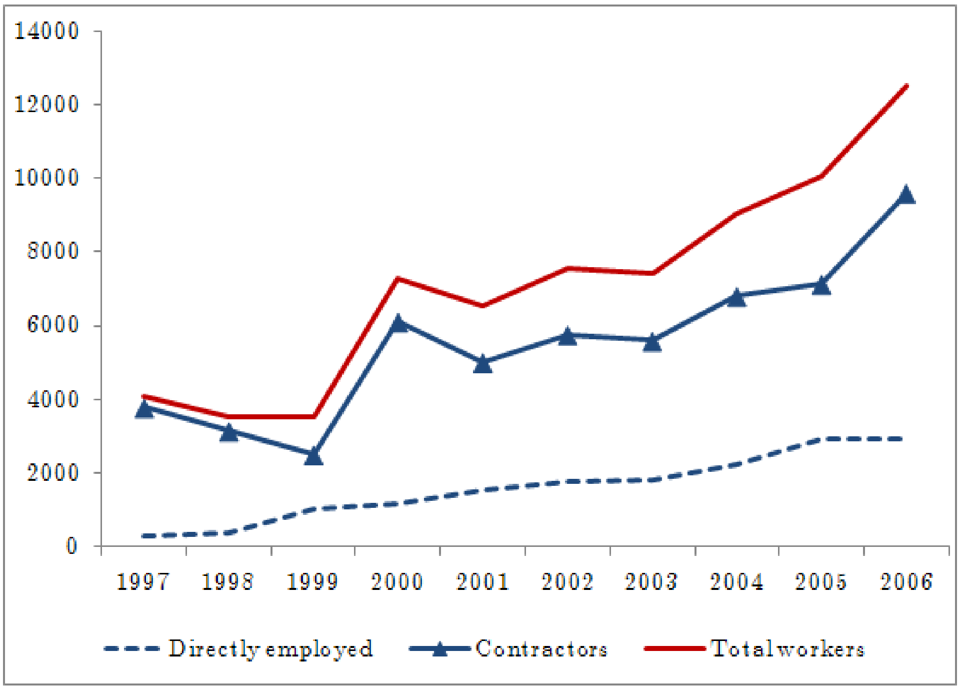

Figure 6: Direct and indirect employment

4 Data and identification strategy

4.1 Data and main variables

Household data The empirical analysis uses data from repeated cross sections of the Peruvian Living Standards Survey, an annual household survey collected by the the National Statistics Office (INEI).[34]

The survey consists of a stratified household sample representative at

[34]The name of the survey is *Encuesta Nacional de Hogares* or ENAHO.

regional level. The regions are defined for statistical purposes and consider both environmental conditions (coast, highlands and forest) and geographical location (north, center and south).[35] We focus on the North Highlands statistical region, where the mine is located. The data set covers 10 years, between 1997 and 2006, and includes more than 7700 households.[36]

Figure 7 shows the area of study and the spatial distribution of the sample. The area covers a surface of almost 40,000 square kilometers and had an estimated population of 1.2 million inhabitants in 2005.[37] The darker shaded areas correspond to districts sampled at least one year, while lighter shaded ones are districts for which there are not available observations. We can see that the sample is evenly distributed over the area of study.

The main purpose of the survey is to measure poverty and living standards. The survey contains detailed information on income, expenditure, socio-demographics (such as gender, age, educational attainment of individuals), composition of the household, housing characteristics like access to public utilities and construction materials, and self-reported incidence of health problems and exposure to crime. The data set also has extensive information on prices and agricultural activity at household level.

Additionally, we obtain the distance from the household's location to the city. We measure distance as the length of the shortest path between the main town of the household's district and Cajamarca city using the existing road network. We performed the calculation using the ArcGIS software and maps produced by the Ministry of Transport of Peru. The road map corresponds to the network available in 2001 and includes only tracks usable by motorized vehicles. In terms of the model, this variable is the empirical counterpart of r and relates to the total transportation cost from one location to the city markets. As suggested by the model, the effect of the mine expansion should decrease with distance to the mine's supplying market. As describe below, we and exploit this heterogenous effects by distance to identify the effect of the mine on the regional economy.

Table 1 shows the estimated mean of the main variables at household

[35] In general, these statistical regions are larger than departments and do not necessarily share the same boundaries.

[36] In line with the model, we restrict attention only to household with an employed household head. This filter reduces the sample by just 46 observations and does not affect the results.

[37] The population estimates are obtained from the ENAHO survey.

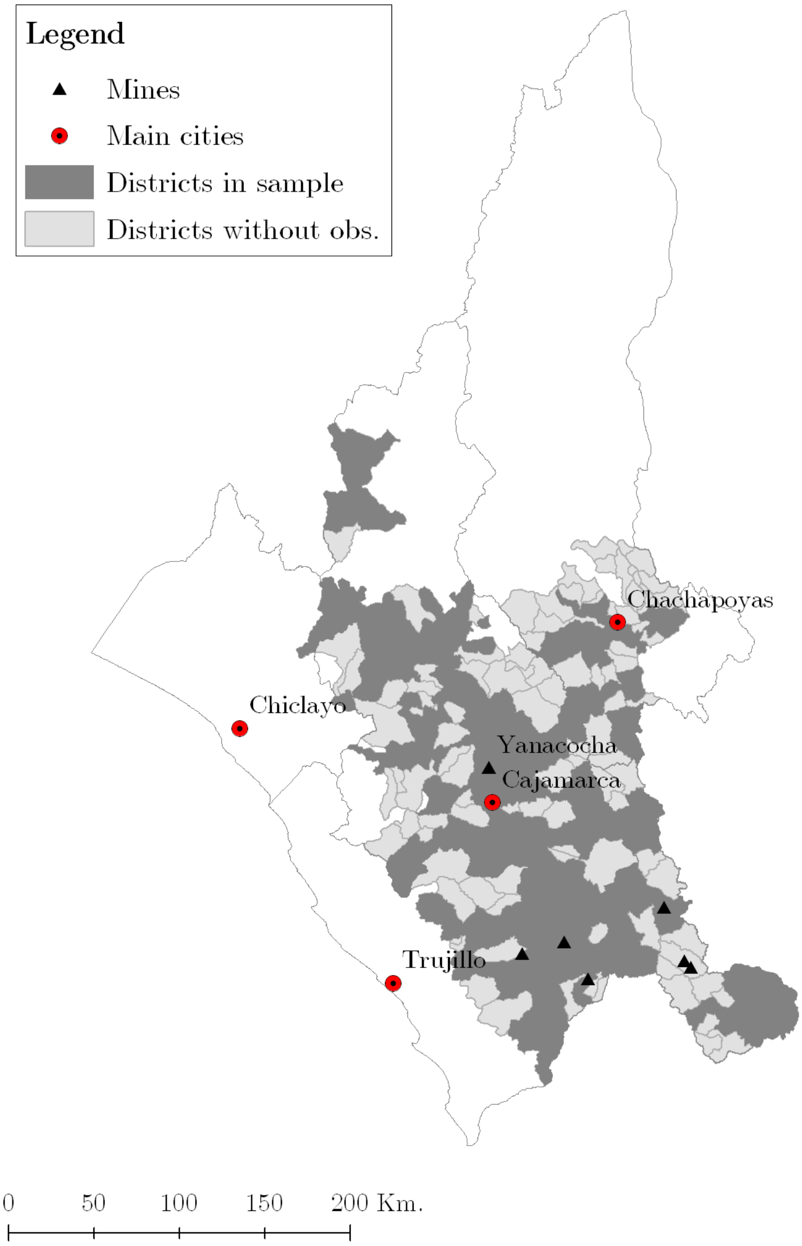

Figure 7: Area of study and spatial distribution of sample

level. Since we are using an stratified sample, the means are estimated using sampling weights. In the second column we report the standard errors of the mean estimates, which are calculated using sampling weights and clustering by primary sampling unit.

Table 1: Mean of main variables at household level

Variables	Mean N=7738	Standard error
Household head		
Years of education	4.1	0.2
Age	48.4	0.7
% female	0.129	0.021
All individuals		
% health problems	0.208	0.012
Household		
Income per capita	143.7	10.8
Expenditure per capita	149.0	8.8
Poverty line	155.9	1.6
% poor	0.651	0.048
% extreme poor	0.374	0.040
% access to electricity	0.163	0.031
% access to water	0.616	0.054
% victim of crime	0.014	0.005
Nr. Household members	4.8	0.2
Nr. Income earners	2.0	0.1
Distance to Cajamarca city (km)	108.2	7.9

Note: The mean and its standard error are calculated using sample weights and clustering by primary sampling unit.

Poverty line as a price deflator The survey provides estimates at household level of nominal income and prices. However, the model provides predictions of the effect of the mine on $y(r)$ and $pa(r)/P(r)$, the real income and relative price of food in a given location. To construct these variables we need a measure of $P(r)$, the general price index in location r.

We use the value of the poverty line as a proxy for $P(r)$. The poverty line is estimated by the Statistics Office as the value of the minimum consumption basket that guarantees an adequate living standard. It is calculated using local prices and varies within region and over time.[38]

A concern with using poverty line as a price index is that it may not be representative of the average consumption basket. Two reasons suggest that this may not be a relevant concern. First, the proportion of households with a consumption below the poverty line is 65%. This implies that the median household is poor and hence the consumption basket used to calculate the the poverty line may not be too different from its actual consumption.

Second, we compare the evolution of the poverty line in Cajamarca city with an index of consumer prices calculated by the National Statistics Office. This price index is only available at city level and it is used to calculate city inflation. Figure 8 depicts both variables, normalized to be equal to 100 in year 1997. Note that the poverty line has a similar trend until 2000 and then grows *faster* than the consumer price index.[39] Thus, if anything, the use of the poverty line as a deflator may lead to an *underestimation* of the actual increase on real income or relative prices.

Firm data To measure the expansion of the mine activities, we collect data from the Yanacocha reports on the mine wage bill, value of local purchases and quantity of gold produced (Minera Yanacocha, 2006). The frequency of this data is annual and covers the period 1993 to 2006.

Local purchases include both goods and services bought to local suppliers and contractors. The wages of workers indirectly employed in the mine, through contractors, are included in this variable. A supplier or contractor is classified as local if it is registered for tax purposes in the department of Cajamarca and the main shareholders are also locals or with permanent residence in the department. The wage bill includes net wages and bonuses paid to workers directly employed by Yanacocha. The data on wages do not separate workers by location; however, almost all the mine's workforce lives in Cajamarca city.

Panel A in Table 2 presents summary statistics for the firm level data over the period 1997 to 2006. The value of wage bill and local purchases

[38]Ideally we would like to construct a price index for each locality; however, we are limited by the lack of representativeness of the sample at more local level.

[39]This pattern is consistent with the model predictions. In particular, recall that the relative price of locally produced food may increase with mine expansion. Then if the poverty consumption basket has a larger proportion of local food than the basket used to calculate the consumer price index, we could expect the value of the poverty line to increase faster.

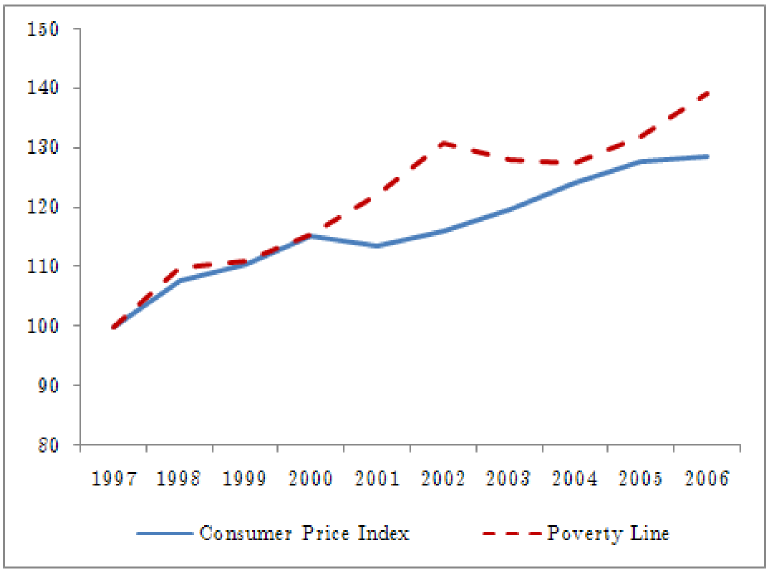

Figure 8: Consumer Price Index and Poverty Line for Cajamarca City

is measured in million of US$ while the quantity produced is measured in million of ounces.

Municipality data We complement the household and firm data with data at municipality level.[40] Municipalities are the lowest tier of autonomous local government with jurisdiction over districts. We use this administrative unit to calculate the distance from the city to the households.

We obtain annual data about revenues and expenditures for each municipality in the North Highlands region and within 400 km from Cajamarca city. This geographical scope corresponds to the distance range observed in the household data. The dataset includes information on 179 municipalities over the period 1998 to 2006.

Our dataset contains detailed information about the sources of revenue, including the amount of mining transfers received. This information provides a reliable measure of the magnitude of the revenue windfall experienced by each local government. The expenditure is also divided between capital and current expenditures. This is a relevant distinction since the mining transfers are earmarked to capital expenditures, and hence we can expect a

[40]We use this data in Section 6.1 to evaluate the role of the fiscal revenue windfall as an alternative explanation of the observed phenomena.

relative increase of this category of expenditure.[41]

Panel B in Table 2 displays some summary statistics. The average municipality has an annual budget of 3.5 million of Nuevos Soles, but a slightly smaller expenditure. The difference is kept by the local government and rolled forward to next periods. in the period of analysis, capital expenditures represented around 40% of total expenditures, while mining transfers were around 12% of total revenue.

Table 2: Summary statistics of firm and municipality data

Variables	Mean	Standard deviation
A. Firm data		
Wage bill	55.5	33.9
Local purchases	42.3	27.7
Gold production	2.18	0.76
% local purchases	0.12	0.05
B. Municipality data		
Total revenue	3.53	8.20
Mining transfer	0.41	2.08
Total expenditure	2.96	6.52
Capital expenditure	1.26	2.42

Note: The value of wage bill and local purchases is measured in million of US$ while the quantity produced is measured in million of ounces. The municipal data is measured in million of Nuevos Soles.

4.2 Identification strategy

The aim of the empirical exercise is to evaluate the effect of the mine expansion on real income, relative prices and measures of welfare.[42] To do that, we exploit two sources of variation. First, we use the significant increase on Yanacocha's wage bill and local purchases started in 2000. As previously mentioned, this growth was driven by the increment on gold extraction and implementation of mine's policies directed at increasing local procurement and employment.

[41] Capital expenditures include mainly the investment on infrastructure projects.

[42] We provide more details about the measures of welfare and the rationale for using them in the next section.

Second, we use distance to the city as a measure of heterogeneous exposure to the mine expansion. As predicted by the model, we expect the effect on real income, if positive, to decrease with distance to the city due to transportation costs. At a large enough distance, the effect should become insignificant because there is no economic interaction between those locations and the city. This prediction suggest that we could use households living in areas farther from the city as a control group of households living in areas closer to it. In the sample, the distance ranges from 0 to 400 km with an average value of around 100 km. We use this threshold to divide the districts in two categories: far and closer to the city.[43]

Our identification strategy is basically a difference in difference, with the expansion of the mine being the treatment and the distance to the city defining treated and control groups. The underlying identification assumption is that the evolution of the outcome variables in areas close and far from the city would have been similar in the absence of the mine expansion.

Figure 9 illustrates the basic idea behind the identification strategy. It plots the conditional mean of real income per capita for households located in districts within 100 km from the city and farther away.[44] Note that until 2001 the real income follows similar trends in both locations. After that, it diverges and increases in areas located closer to the city relative to areas located farther away.

The similarity of trends in both areas before the expansion of the mine is a necessary condition for the validity of our difference in difference strategy. There may be, however, other unobserved time-varying factors correlated with the expansion of the mine and affecting differently areas closer and farther from the city, which would invalidate our identification assumption.

In Section 7, we explore the possibility of confounding factors by including heterogenous non-parametric trends based on observables (such as urbanization, density, type of economic activity or proximity to a city) and by performing a falsification test using other cities in the north of Peru that do not have a nearby mine. In all cases, the results are similar to the baseline regressions and provide suggestive evidence that our estimates capture the effect of the mine expansion.

[43] In Section 7.3 we check the robustness of the results using alternative definitions of proximity and distance.

[44] The mean is conditional on schooling, age and gender of the household head, access to water and electricity, number of household members and number of income earners.

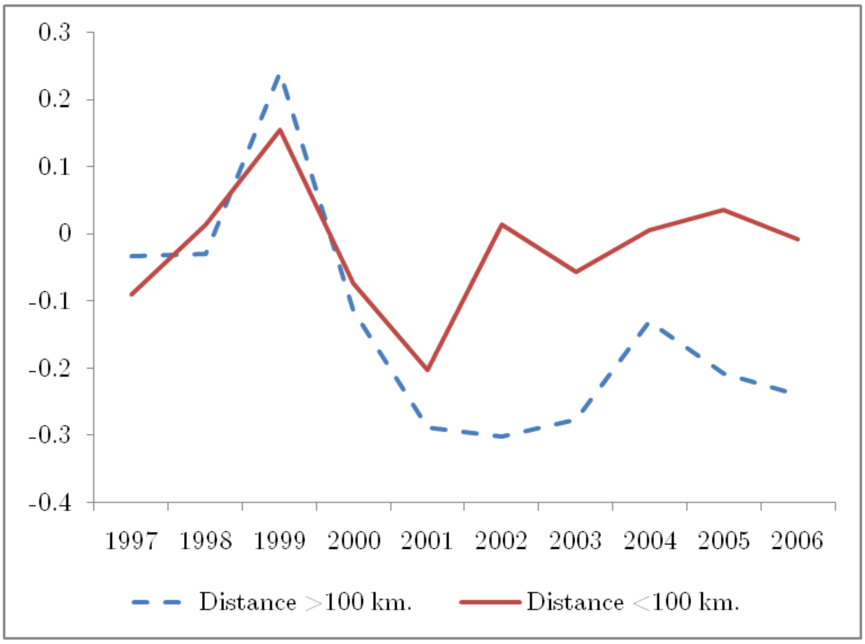

Figure 9: Conditional mean of real income

To formally evaluate the effect of the mine expansion on real income, relative prices and measures of welfare, we estimate the following regression:

$$\ln y_{hdt} = \alpha_d + \eta_t + \beta(\ln M_t \times D_d) + \mathbf{X}_{hdt}\gamma + \varepsilon_{hdt}, \tag{20}$$

where y_{hdt} is the outcome variable of household h in district d in year t. The outcome variables could be real income, relative price or a measure of welfare. M_t is a measure of the mine activity, lagged two periods to allow adjustments in market prices.[45] In the baseline specification we use the value of the mine's wage bill and local purchases, but we also check the robustness of the results using alternative measures such as the value of local purchases or quantity of gold produced. D_d is a dummy equal to 1 if the district where the household lives is within 100 km of Cajamarca city, and 0 otherwise. This dummy identifies households living close and far to the city and thus defines the treated and control group (see Figure 10 for a map of the spatial distribution of both groups). $\mathbf{X}_{hdt}$ is a vector of time-varying control variables. In this specification, the parameter of interest is β which

[45]The results are similar using other time lags.

captures the effect of the mine expansion.

All regressions include year (η_t) and district (α_d) fixed effects to account for common regional shocks and unobserved time-invariant district characteristics. The regressions are estimated using sample weights and clustering the standard errors at the level of the primary sampling unit. The primary sampling unit is a geographical unit smaller than districts and defined for sampling purposes. The estimation procedure takes into account that the sample is not a random draw of the region's population but an stratified sample.[46] The reason for clustering is to account for spatial correlation of households exposed to similar shocks and market conditions, or surveyed simultaneously.

[46]See Magee et al. (1998) for a discussion on the use of sample weights with complex survey data.

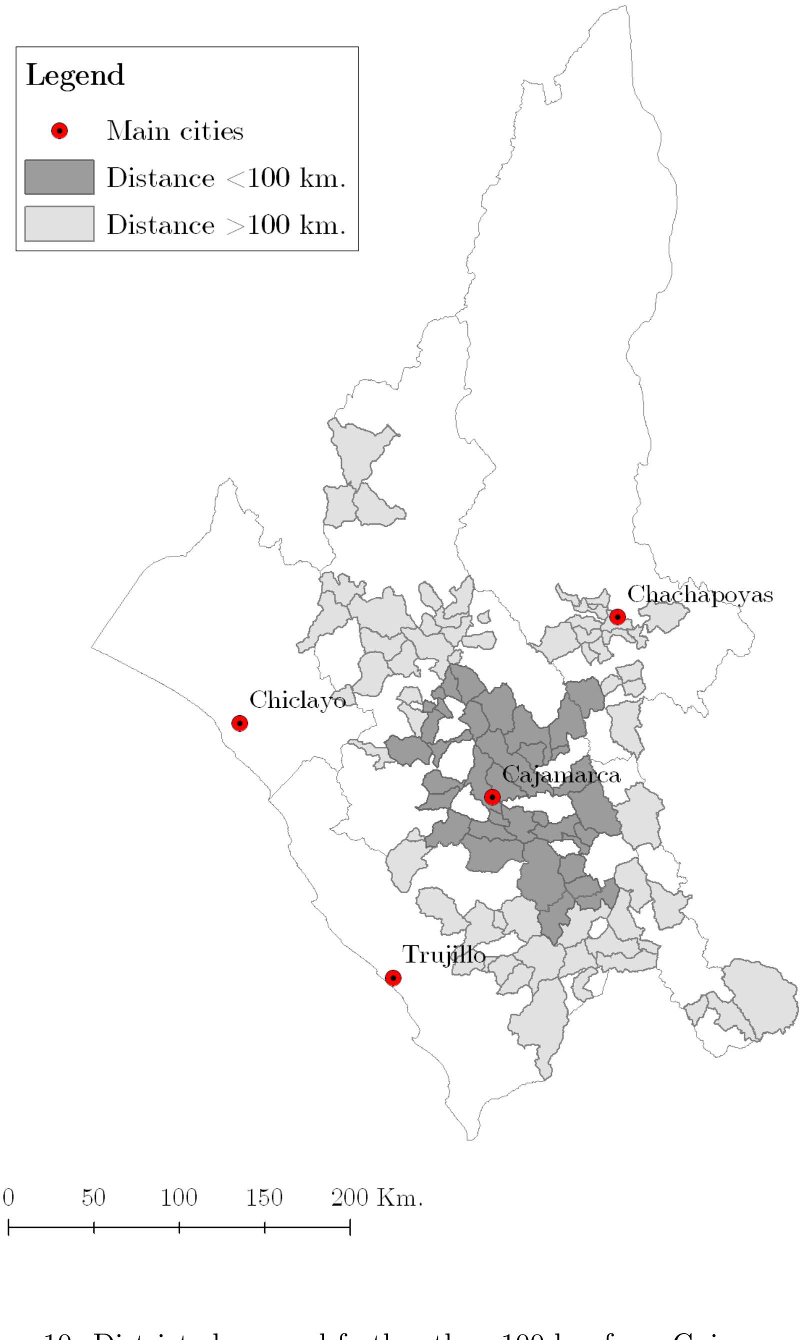

Figure 10: District closer and farther than 100 km from Cajamarca city

5 Main Results

The model highlights a market channel for the extractive industry to affect a regional economy. In this story, the mine expansion increases the demand for local labor. This demand shock increases wages in the city, but it is partially offset by migration of rural workers. The net effect on city income is ambiguous due to the crowding out of other sectors, that may reduce alternative sources of income such as dividends from export firms. However, if city income increases, the price of local food crops also increase due to the lower supply. In turn, this translates into higher income in rural locations, but the effect decreases with distance due to transportation costs. In the absence of externalities, the increase on real income also represents increase on welfare.

This mechanism resembles the classical multiplier effect, with growth in the mining industry creating ripple effects in other sectors of the economy. The magnitude of the effect depends on the extent of backward linkages connecting the extractive industry to the rest of the economy.[47] The benefits accrue to all residents of the regional economy - not only to the ones directly employed in the mine or living in the city - due to mobility of goods and labor, but do not reach households living outside the city hinterland.

In this section we first evaluate the model predictions regarding the effect on real income and relative price of local food crops. Then, we explore the effect on welfare and the re-distributional effects of the mine expansion.

5.1 Effect on Real Income

Columns (1) and (2) of Table 3 show the estimates of β using the baseline specification (20) with different measures of the mine activity. Panel A uses the value of the mine wage bill and local purchases as a measure of mine activity while panels B and C use the value of the mine local purchases and the quantity of gold produced, respectively. All regressions include year and district fixed effects and a vector of control variables. In column (1) we use the full set of control variables including characteristics of the household head such as education, age, gender and dummies indicating the occupation industry and type of job (e.g. independent worker, employee,

[47]The model suggests that if the mine does not buy local inputs, its expansion would not have any effect on the regional economy.

wage worker, etc.). It also includes household characteristics lick access to water, electricity and number of household members and income earners. In column (2) we use a more parsimonious specification with only the household head's schooling, age and occupation industry.

In all the cases, the estimates of β, the change on real income associated to the mine expansion, are positive and significant. The similarity of results shows that the results are not sensitive to different measures of the mine activity or alternative set of control variables.[48] In what follows, we use the mine wage bill and local purchases as our preferred measure of the mine activity because it relates, more closely, to the mine's demand of local inputs.[49]

The positive relation between the mine expansion and real income suggests that the crowding out of local firms (and the subsequent reduction of their profits) does not affect the income of the average household. This is not surprising since most people in the sample are wage workers or farmers, and hence less likely to be affected by reduction on firm profits.[50]

Under the assumption that the evolution of income in locations far and closer to the city would have been similar in the absence of Yanacocha, we can interpret these results as the evidence of a positive effect of the mine on real income. The magnitude of the effect is economically significant: the smallest estimate using the full specification suggests that a 10% increase in the mine's activity is associated to an increase of 1.7% in the real income of households located closer to the city. Note that the evolution of Yanacocha implies large changes in household incomes since, by any measure, the activity of the mine has multiplied by at least a factor of 2.

The model predicts that the effect of the mine expansion on real income decreases with distance. This heterogenous response justifies the use of households far from the city as controls for households closer to the city. To explore this hypothesis further we present two additional specifications.

[48] We also use real expenditure per capita as a dependent variable. The results, not reported, are similar.

[49] Note that the quantity of gold produced can be used as an instrument for the mine wage bill and local purchases. In that case, the results in Panel C would correspond to the reduced form estimates. We run a 2SLS regression using the quantity of gold as an instrument for the mine's wage bill and local purchases, and find an estimate for $\beta = 0.306$, significant at 5%.

[50] In the sample, 95% of individuals work as independent workers, laborers, employees or family workers.

Table 3: Effect of Yanacocha's expansion on real income per capita

	Ln(real income per capita)			
	(1)	(2)	(3)	(4)
A. Measure of mine activity = mine's wage bill and local purchases				
Mine activity	0.174**	0.161*		
× distance < 100 km	(0.078)	(0.088)		
Mine activity			-0.128**	-0.102
× continuous distance			(0.062)	(0.064)
R-squared	0.521	0.349	0.521	0.348
B. Measure of mine activity = mine's local purchases				
Mine activity	0.215**	0.188*		
× distance < 100 km	(0.091)	(0.101)		
Mine activity			-0.151**	-0.119*
× continuous distance			(0.060)	(0.064)
R-squared	0.523	0.35	0.522	0.349
C. Measure of mine activity = mine's gold production				
Mine activity	0.313**	0.316**		
× distance < 100 km	(0.131)	(0.154)		
Mine activity			-0.207**	-0.185*
× continuous distance			(0.100)	(0.106)
R-squared	0.523	0.35	0.522	0.349
Observations	7738	7738	7738	7738
Full set of controls	yes	no	yes	no

Notes: Robust standard errors in parentheses. Standard errors are adjusted for clustering at primary sampling unit level. * denotes significant at 10%, ** significant at 5% and *** significant at 1%. All regressions include year and district fixed effects. The full set of control variables includes: household head's education, age, gender and dummies indicating her industry of occupation and type of job, plus household access to water, electricity and number of household members and income earners. Columns (2) and (4) use a smaller set of controls: household heads' schooling, age and dummies of her industry of occupation. Columns (3) and (4) use a measure of continuous distance, expressed in hundreds of kilometers.

First, we use a flexible specification for the measure of distance. In particular, we divide the households in six groups according to the distance to the city and create dummies corresponding to each category. The categories start with households living in Cajamarca city and then group them in blocks of up to 50 km, with the last category containing all households located at least 200 km from the city. Then we estimate the baseline regression (20) with the full set of controls, using the mine wage bill and local purchases as measure of the mine activity interacted with each of the distance dummies. The omitted category is the group of households living farther than 200 km.

Figure 11 shows the estimates of β for households located at each of the distance brackets, as well as the 95% confidence interval. The estimates are positive and significant for households located within 100 km of Cajamarca city, but become insignificant for households located in farther locations. These results provide the basis for using a dummy variable of distance and reduces concerns about the observed average effect being driven exclusively by city residents.[51]

Second, we run a regression on the interaction of the measure of the mine's activity and the *continuous* measure of distance, expressed in hundreds of kilometers. The results are displayed in columns (3) and (4) of Table 3. Note that in almost all the cases, the estimated parameter is negative and significant. Taken together, these results are consistent with the prediction of the model that the effects decrease with distance.[52]

5.2 Effect on Relative Prices

The model predicts an increase on the relative price of locally produced goods - such as foodstuff - as a consequence of the demand shock initiated by the mine activities. This price increment is responsible for transmitting the effect among rural agricultural producers, who do not participate directly in the city's labor market. In contrast, the model is mute regarding the effect on the relative price of imported goods. Their relative price may increase or decrease (see Proposition 3).

To test these predictions, we start by identifying the main crops in the

[51] For example if the distribution of household income improvements was only concentrated in Cajamarca city, the reduced form estimates would be just averaging out large positive effects in the city and negative or zero effects in the vicinity.

[52] In section 7.3 we evaluate the robustness of the results to the use of alternative measures of distance and non-linearities.

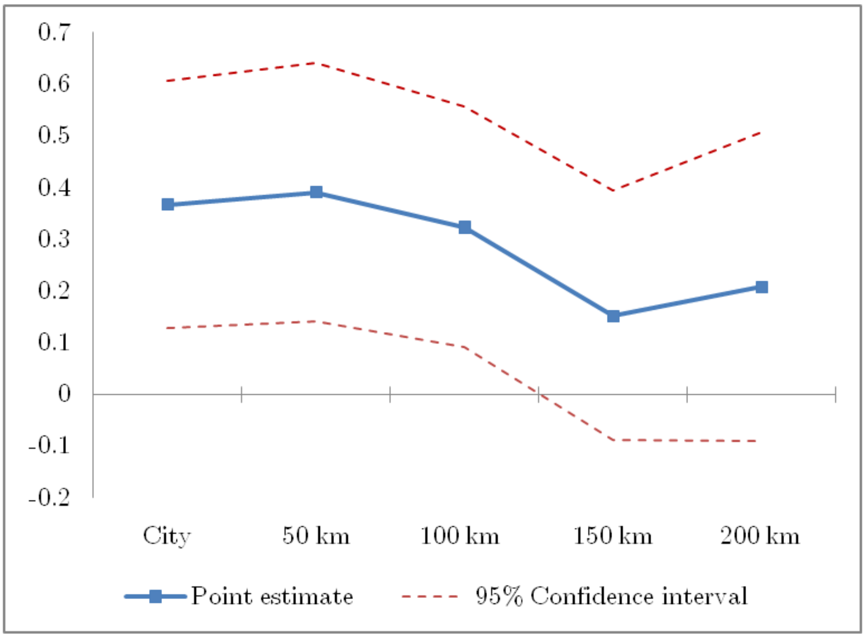

Figure 11: Effect of Yanacocha's wage bill and local purchases on household income, by distance to Cajamarca city

area of study. We use information from the household survey about agricultural production and rank the crops according to its contribution to the regional agricultural gross product.[53] In our sample, the two most important crops are potatoes and maize. Together they account for almost half of the agricultural gross product.[54] These results are consistent with the data from the 1994 Agricultural Census. According to the census, potatoes and maize were the most widespread crops in the highlands of Cajamarca representing more than 50% of the land cultivated with annual crops.

For each crop we obtain the farm gate price received by the producer based on reported quantity and value of production. In addition, we calculate the unit value paid by consumers using information about total expenditure and quantity purchased. These variables do not correspond exactly to market prices because they are also affected by the household's quality choice. The producer prices and unit values are then divided by the value

[53]The survey has detailed information about the production and value of individual crops at producer level.

[54]In the period 1997 to 2006, they represented 30% and 16% of the value of agricultural production, respectively. Their contribution remained relatively constant over the period of analysis.

of the poverty line to obtain measures of relative prices.

To explore the effect of the mine expansion, we estimate the baseline regression (20) using as dependent variable the logarithm of the measure of relative prices. When using unit values, we include as an additional control variable the logarithm of the household real income to account for quality choices (Deaton, 1997).[55] In the case of prices at producer level, we do not include any additional control. The reason is that the inclusion of income or expenditure as controls would bias the results since the income of agricultural producers also depends on crop prices.

Table 4 presents the estimates using the measures of prices at producer and consumer level. In most cases, the estimates suggest that the relative prices of local crops in areas closer to the city increase relative to prices in markets located in farther locations. The use of year fixed effects reduce concerns that this relation is driven by an underlying common trend. The lack of effect in the case of the price of maize received by producers (Column 2) may be due to the poor development of the maize market. According to the survey, only 40% of the maize production was sold to the market. In contrast, the proportion sold to the market of the potatoes production was 63%. Note, however, that there is a positive and significant effect on the price paid by consumers.

As a second step, we replicate the same empirical exercise using unit values of commodities consumed by a large number of households and less likely to be locally produced such as rice, sugar, cooking oil and canned fish.[56] In the model, these goods correspond to the import good. Similar to the previous regressions, we use the baseline specification 20 and include the log of real income per capita as an additional control variable. The results are shown in Table 5. In contrast to the case of local crops, the effect of the mine on relative prices is negative or insignificant. This result reduces concerns that the observed increase on relative price of locally produced food is driven by a general increase of food stuff.

We can go further and explore the changes on income relative to the price of local food. As predicted by the model, we expect the increase on

[55]The underlying assumption is that quality is correlated with income.

[56]For example, rice represents only 1.6% of the agricultural product in the sample, while sugar and cooking oil are processed food stuff traded at national level. Of particular interest is canned fished (sardine), because the region is landlocked and thus the production is certainly not local.

Table 4: Effect of mine expansion on price of local goods

	Ln relative price of:			
	Potatoes		Maize	
	(1)	(2)	(3)	(4)
Mine's wage bill and local purchases × distance < 100 km	0.128*	0.084*	-0.026	0.110***
	(0.074)	(0.050)	(0.082)	(0.038)
Prices reported by:	producer	consumer	producer	consumer
Real income per capita	no	yes	no	yes
Observations	3407	4072	4253	4000
R-squared	0.417	0.524	0.374	0.19

Notes: Robust standard errors in parentheses. Standard errors are adjusted for clustering at primary sampling unit level. * denotes significant at 10%, ** significant at 5% and *** significant at 1%. All regressions include year and district fixed effects. Columns (1) and (3) do not include any additional control variable. Columns (2) and (4) include ln(real income per capita) as control variable.

Table 5: Effect of mine expansion on price of non-local goods

	Ln relative price of:			
	Rice	Sugar	Cooking oil	Canned fish
	(1)	(2)	(3)	(4)
Mine's wage bill and local purchases × distance < 100 km	0.021	0.037	0.020	-0.111*
	(0.033)	(0.029)	(0.028)	(0.064)
Prices reported by:	consumer	consumer	consumer	consumer
Real income per capita	yes	yes	yes	yes
Observations	5174	5520	2349	925
R-squared	0.416	0.349	0.591	0.486

Notes: Robust standard errors in parentheses. Standard errors are adjusted for clustering at primary sampling unit level. * denotes significant at 10%, ** significant at 5% and *** significant at 1%. All regressions include year and district fixed effects, and ln(real income per capita) as control variable.

income deflated using the price of local crops to be smaller than the increase on income deflated by a more general price index (Proposition 4). The main reason is that the increase on rural income is mostly driven by higher food prices, hence the effect of the mine expansion on income deflated by food price should be negligible.

To test this prediction we re-estimate the baseline regression (20), using alternative measures of prices to deflate the nominal income. In particular, we use the unit value of potatoes and maize. In both cases, we take the average value by primary sampling unit to reduce measurement errors and obtain better proxies of the underlying prices.

Table 6 displays the results. Columns (1) and (2) use as explained variable the log of the income divided by the unit value of potatoes or maize. Column (3) displays the baseline result using the poverty line as deflator. Note that the estimated effect of the mine expansion is positive but insignificant when using the unit values of local crops as deflators. The point estimates are also smaller than the one obtained using poverty line as a deflator.

Table 6: Effect of mine expansion on real income, using alternative price deflators

	Ln(relative income per capita)		
	(1)	(2)	(3)
Mine's wage bill and local purchases × distance <100 km	0.137 (0.091)	0.115 (0.091)	0.174** (0.078)
Deflator of nominal income	unit value of potatoes	unit value of maize	poverty line
Observations	7132	7089	7738
R-squared	0.562	0.525	0.521

Notes: Robust standard errors in parentheses. Standard errors are adjusted for clustering at primary sampling unit level. * denotes significant at 10%, ** significant at 5% and *** significant at 1%. All regressions include year and district fixed effects, and control variables.

5.3 Effect on Welfare

So far we have shown that the mine has had a positive effect on real income. In the context of the model, this increase of real income translates into higher household utility. However, in the presence of negative externalities this is not necessarily the case. The negative effects of the mine may offset the benefits of having a higher income and leave households neither better or worse off. In that case, higher wages would just compensate for the mine negative externalities.

For example, the mine may have increased the actual, or expected, levels of pollution or environmental degradation with negative effects on health. Similarly, the influx of new population and growth of the city may be linked to increase on criminal activity or anti-social behavior. There is anecdotal evidence of concerns among the population about the risk of water pollution due to the use of highly toxic chemicals in the mine, and discomfort associated to the perceived increase on prostitution and crime (Pascó-Font et al., 2001, p. 156).

To evaluate the effect on welfare, we proceed in the following way. We first treat the mine as an "amenity" and apply the analytical framework developed by Roback (1982) that suggests using changes on housing prices as a measure of net welfare effects. Then, we evaluate changes on indicators of negative spillovers - such as incidence of health problems and crime.

5.3.1 The Mine as an Amenity

Using the simplest version of the model in Roback (1982), the mine can be studied as an amenity that generates locational externalities (positive or negative) in a region with perfect labor mobility and a fixed supply of land. If the mine generates negative externalities to workers (e.g. pollution), there is an outflow of people from the region. This emigration increases wages but *decreases* land rents. Alternatively, if the mine creates positive externalities to producers (e.g. due to agglomeration economies) there is an increase on labor demand and wages.[57] The higher wages attract immigrant workers and *increase* the rental price of land. If both forces are at play, wage unequivocally increases but the effect on rent is ambiguous and depends of the relative magnitude of both externalities.

[57]This effect is similar to the one emphasized in our model in Section 2

A particularly useful implication of this model is that changes on property values reflect the net effects on welfare (for an empirical application see Greenstone and Moretti (2003)). In particular, lower land rents would suggest that the increase on real income is just compensating workers for negative externalities. In contrast, an increase on rents would provide evidence of a net positive effect.

To evaluate this hypothesis, we use data from the household survey on self-reported house rents. This variable was obtained asking home owners about the minimum price they would require for renting their property.[58] We calculate a proxy for relative land rents dividing the reported house rents by the value of the poverty line.

Then, we estimate the baseline regression (20) with the logarithm of the relative house rent as dependant variable. In addition to year and fixed effects, the regression includes controls for observable house characteristics that may affect the property value such as type of urban settlement, construction materials (walls and floor), number of rooms and access to utilities (water, sewage, electricity and phone landline). In addition, we also include household socioeconomic variables - like schooling, age and gender of the household head and number of members and income earners - to account for systematic biases in the report of rental price. This empirical specification corresponds to the hedonic regression, a widely used method to assess the value of environmental amenities.[59]

Table 7 displays the results using different sets of control variables. Column (1) includes only house characteristics while column (2) runs the complete specification with socio-economic controls. In both cases, we find evidence of a positive relation between the mine activity and house rents. The magnitude of the effect is significant, with an estimated elasticity of 0.24. This evidence is consistent with the mine activities having a positive net effect on household's welfare.

[58] We did no use actual rent prices due to the small number of observations. In the sample, only 6% of households live in rented accommodation.

[59] See Black (1999) and Greenstone and Gallagher (2008) for examples of empirical applications.

Table 7: Effect of mine expansion on house rents

	Ln (relative value of house rent)	
	(1)	(2)
Mine's wage bill and local purchases × distance < 100 km	0.250**	0.238**
	(0.095)	(0.092)
House characteristics	yes	yes
Socio-economic controls	no	yes
Observations	7699	7696
R-squared	0.635	0.650

Notes: Robust standard errors in parentheses. Standard errors are adjusted for clustering at primary sampling unit level. * denotes significant at 10%, ** significant at 5% and *** significant at 1%. All regressions include year and district fixed effects. Column (1) include as additional controls only house characteristics such as type of urban settlement, construction materials (walls and floor), number of rooms and access to utilities (water, sewage, electricity and phone landline). Column (2) adds socio-economic characteristics of the household such as schooling, age and gender of the household head and number of household members and income earners.

5.3.2 Health and Crime

A main limitation of the previous analysis is that the self-reported rental prices may fail to fully capture the net effect on household welfare. This may happen, for example, if housing markets are incomplete, if self-reported rental prices are biased or if individuals lack information regarding the negative effects associated to the mine.

To address this issue, we test directly for the presence of some negative effects associated to the expansion of the mine. In particular, we use data from the household survey to construct dummy variables to indicate whether an individual had a health problem, and whether a household member was victim of a criminal activity.[60] Then, we use these dummies as the dependent variables in the baseline regression (20) and estimate it using a linear probability model. As control variables we use an indicator of whether the household lives in a urban or rural area, access to water, sanitation and electricity, number of household members and income earners, and individual's sex and age.[61]

Table 8 shows the results. Column (1) uses the whole sample of individuals, including children, while in column (2) we restrict the sample to children under the age of five, who may be more vulnerable to negative health spillovers. Note that in both cases the incidence of self-reported health problems has actually decreased with the expansion of the mine. Column (3) shows that there is no apparent increase in crime associated with the expansion of the mine.

Nonetheless, we need to interpret these results with caution. They only suggest that there is no evidence that individuals in the area of influence of the mine have suffered more occasional illnesses, that could result from a more polluted environment. But, we cannot say anything about long run effects, such as general deterioration in health or chronic afflictions that could result from exposure to the activities of the mine. Similarly, the measure of crime only informs us about the perceived level of crime but may fail to account for other forms of social disorder or crimes within the household. Moreover, it may reflect changing perceptions of crime.

---

[60] The survey questions are "In the last four months, have you felt sick, suffered a chronic disease or an accident?" and "In the last 12 months, has any member of the household been affected by a criminal act?".

[61] The results are robust to the exclusion of the control variables.

Taken together, these results do not rule out the presence of negative externalities. They suggest, however, that their magnitude, at least as perceived by the residents, may not be too important. This finding is consistent with the observed net improvements on household welfare discussed previously.

Table 8: Yanacocha's expansion and measures of health and crime

	Health problems		Crime
	(1)	(2)	(3)
Mine's wage bill and local purchases × distance < 100 km	-0.087***	-0.008	0.001
	(0.039)	(0.057)	(0.008)
Sample	All individuals	Children under 5 years	Households
Observations	39674	4189	6663
R-squared	0.076	0.157	0.048

Notes: Robust standard errors in parentheses. Standard errors are adjusted for clustering at primary sampling unit level. * denotes significant at 10%, ** significant at 5% and *** significant at 1%. *Health problems* is a self-reported measure of whether an individual was sick in the recent past, while *Crime* is a self-reported measure of anyone in the household being victim of a crime in the recent past. All regressions are estimated using a linear probability model. Columns (1) and (2) include as control variables an indicator of the house being in a urban area, access to water, sanitation and electricity, number of household members and income earners, and individual's sex and age. Column (3) use similar controls but exclude individuals' age and sex.

5.4 Distributional Impacts

The previous results are only informative of the effect of the mine expansion on the income, and welfare, of the average household. They do not tell us, however, whether the operations of the mine have reduced poverty or whether the effects were distributed evenly across different income groups. The increase on average income may hide negative re-distributional effects with only the richer benefitting from the mine expansion. Even if equally distributed, the income increase may have been insufficient to close the poverty gap.

We address these questions in two steps. First, we evaluate the effect of the mine expansion on measures of poverty. To do so, we estimate the baseline regression (20) using as dependent variables dummies indicating whether a household is poor or extreme poor. The difference between these two categories is the definition of the minimum consumption basket, which is more restrictive for the case of extreme poor.[62] We estimate the regression using a linear probability model and interpret the coefficients as the change on the probability of being poor or extreme poor. Table 9 displays the results. Note that the mine expansion is associated to reductions on the probability of being poor or extreme poor.

Table 9: Yanacocha's expansion and poverty

	Poor (1)	Extreme poor (2)
Mine's wage bill and local purchases × distance < 100 km	-0.107*** (0.036)	-0.075** (0.056)
Observations	7738	7738
R-squared	0.406	0.3572

Notes: Robust standard errors in parentheses. Standard errors are adjusted for clustering at primary sampling unit level. * denotes significant at 10%, ** significant at 5% and *** significant at 1%. All regressions are estimated using a linear probability model.

Second, we explore the distributional impact of the mine expansion estimating the the effect of the mine on real income at different points of the conditional income distribution using quantile regressions.[63] The model specification is the same as the baseline regression (20), but the estimator does not use the sampling weights nor cluster the errors by primary sampling unit.

Figure 12 plots the estimates of β - the effect of the mine expansion on real income - (on the vertical axis) and the confidence interval at 95% at different quantile values. The estimated parameter is positive and significant

[62] In the sample, 65.2% of households are poor, while 37.4% are classified as extreme poor.

[63] See Koenker and Hallock (2001) for a survey of the literature on quantile regressions.

for all quantiles except for the top 20%. These results mean that households with income below the 80% percentile, conditional on the control variables, experienced an increase of real income associated to the mine expansion in areas within 100 km from Cajamarca city. We cannot conclude from this result that inequality has not increased, but we can claim that the positive effects were evenly distributed across income groups, even among the poorer households.

In this section we have provided evidence that the increase in income that followed the expansion of local purchases by Yanacocha is also associated with a positive net welfare effect. There is no evidence of short run negative externalities, poverty has decreased and even the poorest of the poor in Cajamarca city and surroundings have obtained an increase in income relative to households in other surrounding areas.

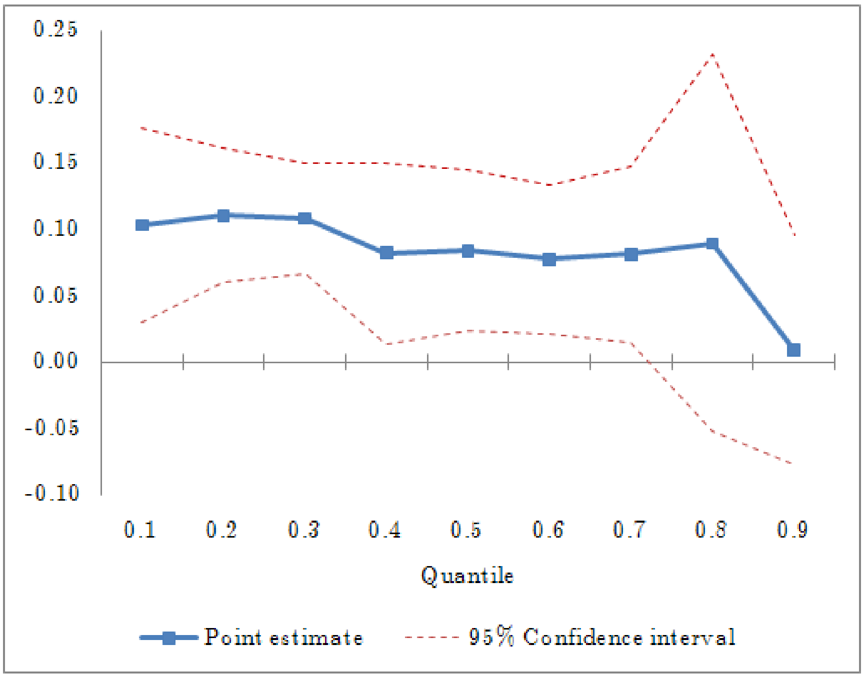

Figure 12: Effect of Yanacocha's wage bill and local purchases on household income, by quantile

6 Alternative explanations

In addition to local purchases and employment, there are at least two other channels through which Yanacocha's activities might have affected the regional economy: a revenue windfall to local governments and direct benefits to local groups, through higher wages or local development projects.

6.1 Fiscal Revenue Windfall

Local governments receive a transfer from the central government, funded with the taxes paid by the mine. In particular, 50% of the corporate tax paid by the mine is distributed to local governments in the region.[64] This mining transfer (called *canon minero*) is allocated according to a formula established by law, that takes into account location of the mine pits, population, density and poverty incidence.[65]

This source of local revenue grew in the last years following the expansion of the mine operations and represented a substantial revenue windfall for local governments. For example, between 1998 and 2006 the total amount of mining transfers in the area multiplied by a factor of 7 and its contribution to the municipal budget increased from 8% to 25%.

This revenue windfall may explain the observed relation between the mine expansion and real income. For example, the additional revenue could have increased public spending and demand of local inputs. Similarly, better public good provision could have enhanced household welfare, and contributed to the increase in housing rents.

In order to understand the mechanism driving the results, we evaluate whether the effects on real income and welfare are driven by a fiscal channel (additional public spending) or by a market channel (mine's demand of local inputs). We address this question in two steps. First, we evaluate the effect of the revenue windfall on municipal total revenue and spending. Second, we evaluate the effect of real income and welfare using both the measure

[64] Mining firms also pay fees for operational rights (derecho de vigencia) which is distributed to the local government where the mine pits are. The total fee is proportional to the extension of the mining operation, regardless of actual production. Its magnitude is much smaller than the *canon minero*.

[65] More specifically, 10% of the tax collected goes to the district where the mine is located, 25% goes to other districts in the province, 40% to the province government and the remaining 25% to the departmental government (Minera Yanacocha, 2006).

of the mine local demand and the revenue windfall received by the local government where the household resides.

6.1.1 Effect on municipal revenue and spending

As a first step, we check whether the revenue windfall from the mining transfer translated into higher revenue and expenditure for local governments. To do that, we estimate the following regression:

$$y_{it} = \alpha_i + \eta_t + \beta x_{it} + \epsilon_{it}, \tag{21}$$

where y_{it} is a revenue or expenditure outcome of municipality i in year t, such as the logarithm of total revenue or the share of capital expenditure while x_{it} is the logarithm of the amount of mining transfer received. This specification includes year and municipality fixed effects, and exploits within municipality variation. The standard errors are clustered at municipality level to address possible serial autocorrelation.

Table 10 displays the results using different dependent variables. All the variables, except the share of capital expenditures, are expressed as logarithms. In all the cases the estimates of β are positive and significant. This result supports the claim that the revenue windfall associated to the mining transfer has increased both the available budget and spending. Capital expenditure increased even faster that other expenditures, as the increase of its share in total expenditure indicates. This result is expected since, by law, the revenues from the mining transfer should be used only for capital expenditures.

6.1.2 Effect on real income and welfare

As a second step, we re-estimate the baseline results on the effect on real income and relative rental prices including, as an additional control, the logarithm of the municipal revenue or expenditure of the district where the household resides. To address possible heterogenous impact of public spending by location, we also estimate the baseline regressions including the measure of municipal revenue or expenditure *interacted* with our measure of proximity to the city (an indicator of the household living less than 100 km to the city or not). If the effect was driven by the revenue windfall, and

Table 10: Effect of mining transfers on municipal revenue and expenditure

	Total revenue	Total expenditure	Capital expenditure	Share of capital expenditure
	(1)	(2)	(3)	(4)
Mining transfer	0.034***	0.030***	0.047***	0.007***
	(0.004)	(0.003)	(0.005)	(0.001)
Observations	1414	1414	1414	1414
Nr. Municipalities	179	179	179	179
R-squared	0.738	0.733	0.476	0.156

Notes: Robust standard errors in parentheses. * denotes significant at 10%, ** significant at 5% and *** significant at 1%. All regressions include year and district fixed effects. All the variables, except the share of capital expenditure, are expressed as logarithms.

subsequent public spending, we could expect the parameter associated to the effect of the mine expansion to become insignificant.

Tables 11 and 12 report the results using the logarithms of real income and relative house rents as dependant variables, respectively. These regressions explore the effect of the mine expansion on real income and household welfare and use the same set of control variables as in the baseline results (see notes in Tables 3 and 7 for further details). In both tables, the regressions in panel A include as additional control variable a measure of municipal revenue or expenditure, while Panel B includes this variable interacted with the dummy of proximity to the city. Columns (1) and (2) use two measures of municipal revenue: the amount received as mining transfer and the total revenue. Columns (3) and (4) uses instead measures of total expenditure and capital expenditures. The difference in the sample size respect to the baseline regressions is due to the lack of observations on municipal budgets for 1997.

Note that in all cases the effect of the mine expansion on real income and house rents remains positive and significant. The estimates are similar in magnitude to the results without municipality control variables. In contrast, the effect of municipality's revenue or expenditure is insignificant or even negative.[66] These results reduce concerns that the observed effect is driven

[66]We also replicate this exercise dropping the measure of mine expansion interacted

by the revenue windfall to local governments. Moreover, they suggest that the market mechanism may have been more effective on enhancing household income and welfare than public spending.

The lack of a positive effect of public spending is surprising. A first explanation is the need of a longer period for public projects to mature. We replicate the results lagging the variables 1 and 2 periods, but the effect remains insignificant.[67] An alternative explanation is that public spending increased well being (through better public good provision) but did not affect income. However, the lack of positive effect on house rents does not support this argument. Finally, it could be that the additional public spending had very small social returns.[68] A similar phenomenon is reported by Caselli and Michaels (2009) in the context of oil-rich Brazilian municipalities. They find that municipalities increased significantly their expenditure using the revenue windfall associated to the oil operations. However they find no evidence of a positive effect on local income. They interpret this result as suggestive that the additional revenue was wasted if not stolen.

6.2 Direct Beneficiaries

There are three groups of individuals that might have benefited directly from the mine expansion: residents of communities located around the mine pits, mine workers, and public sector workers.

In the first case, inhabitants of local communities may have benefited from the development projects started and financed by Yanacocha. These projects are aimed to rural communities adjacent to the mine pits and focus mostly on construction and staffing of schools and health centers, provision of electricity and sanitation, and maintenance of local transport infrastructure.[69] Similar to the expansion of mining transfers, the expenditure on local development projects increased significantly in the period of analysis

with distance and keeping only the municipal variables. The results, not reported, are similar.

[67]We cannot explore longer lags since we only have budgetary data from 1998. Additionally, the use of lags reduces the number of observations in the period before the mine expansion and may attenuate the results.

[68]Anecdotal evidence suggests that some recipients of mining transfers embarked in unproductive projects such as refurbishing the town main square or erecting monuments.

[69]Yanacocha considers as potential beneficiaries the inhabitants of 65 villages adjacent to the mine pits with a population of around 28,000 inhabitants (Minera Yanacocha, 2002, p. 24).

Table 11: Effect of the mine on real income, controlling for municipal revenue or expenditure

	Ln(real income per capita)			
	(1)	(2)	(3)	(4)
Panel A: Adding municipal revenue or expenditure				
Mine's wage bill and local purchases × distance < 100 km	0.175**	0.180*	0.164*	0.164*
	(0.083)	(0.094)	(0.089)	(0.089)
Municipal revenue or expenditure	-0.025	-0.142	0.045	0.045
	(0.209)	(0.093)	(0.126)	(0.126)
Observations	7738	6305	6305	6305
R-squared	0.521	0.536	0.536	0.536
Panel B: Adding municipal revenue or expenditure × proximity to city				
Mine's wage bill and local purchases × distance < 100 km	0.250*	0.295*	0.361**	0.169
	(0.129)	(0.172)	(0.152)	(0.123)
Municipal revenue or expenditure × distance < 100 km	-0.081	-0.178	-0.346*	-0.015
	(0.056)	(0.184)	(0.185)	(0.144)
Observations	7138	6305	6305	6305
R-squared	0.524	0.536	0.536	0.536
Municipal revenue or expenditure	Mining transfers	Total revenue	Total expenditure	Capital expenditure

Notes: Robust standard errors in parentheses. Standard errors are adjusted for clustering at primary sampling unit level. * denotes significant at 10%, ** significant at 5% and *** significant at 1%. All regressions include year and district fixed effects and the same control variables of the baseline regressions (see notes of Table 3). The measures of municipal expenditure or revenue and mine wage bill and local purchases are expressed as logarithms of the total amount.

Table 12: Effect of the mine on house rents, controlling for municipal revenue or expenditure

	Ln (relative value of house rent)			
	(1)	(2)	(3)	(4)
Panel A: Adding municipal expenditure or revenue				
Mine's wage bill and local purchases × distance < 100 km	0.239**	0.360***	0.353***	0.353***
	(0.096)	(0.102)	(0.101)	(0.101)
Municipal revenue or expenditure	-0.016	-0.230**	-0.194*	-0.194*
	(0.242)	(0.097)	(0.106)	(0.106)
Observations	7696	6365	6365	6365
R-squared	0.650	0.667	0.666	0.666
Panel B: Adding municipal expenditure or revenue × proximity to city				
Mine's wage bill and local purchases × distance < 100 km	0.286**	0.472**	0.447**	0.381**
	(0.131)	(0.206)	(0.176)	(0.159)
Municipal revenue or expenditure × distance < 100 km	-0.055	-0.182	-0.171	-0.059
	(0.056)	(0.218)	(0.191)	(0.164)
Observations	7696	6365	6365	6365
R-squared	0.650	0.667	0.667	0.667
Municipal revenue or expenditure	Mining transfers	Total revenue	Total expenditure	Capital expenditure

Notes: Robust standard errors in parentheses. Standard errors are adjusted for clustering at primary sampling unit level. * denotes significant at 10%, ** significant at 5% and *** significant at 1%. All regressions include year and district fixed effects and the same control variables of the baseline regressions (see notes of Table 7). The measures of municipal expenditure or revenue and mine wage bill and local purchases are expressed as logarithms of the total amount.

from around US$ 4 million in 1997 to US$ 34 million in 2006.

In the case of mine workers, the benefit may be explained by the higher wages paid by the mine. For example, in 1997 the average salary for a Yanacocha employee was almost three times the salary for a similar job in Cajamarca city (Pascó-Font et al., 2001, p.165).[70] In the case of public workers, their income would have increased if the revenue windfall associated to mining transfers was redistributed in the form of higher wages or in-kind benefits.

A relevant concern is that the observed effect on the average real income may be just driven by the positive effect of local development projects or higher wages received by mine workers or public servants, not by the multiplier effect highlighted in the model.

To check this possibility, we use data on the industry of occupation reported by employed individuals and identify workers in mining and public service.[71] We classify education workers as part of the public sector given that most teachers are in the government payroll. Additionally, we identify households living in the same districts where the towns targeted by Yanacocha's development projects are located. We consider these groups to be a broader definition of potential direct beneficiaries of the mine activities. In our sample, these households represent around 15% of total observations. Then we re-estimate the regressions of the effect on real income and welfare excluding these three groups of households.

Table 13 displays the results. Panel A uses real income as dependant variable, while panel B uses the relative house rental price,. Column (1) presents the baseline results with the whole sample. Columns (2) to (4) gradually reduce the sample excluding public sector workers, mining workers and households living adjacent to the mine. Note that in all the cases the results are similar to the obtained with the whole sample. This evidence suggests that the observed effects on income and welfare are not driven by groups that may receive direct benefits from the mine, and supports the claim that the main driver is the multiplier effect from the mine's demand shock.[72]

[70] The higher wages may be a compensation for the (potentially) more dangerous working environment or efficiency wages paid for roles deemed as vital for the mine's normal operation.

[71] The survey reports the 2-digit International Standard Industry Code of the main activity of working individuals.

[72] We are unable to evaluate in a reliable manner the effect among direct beneficiaries

Table 13: Effect on households not directly employed by public sector, mining industry or living adjacent to the mine

	(1)	(2)	(3)	(4)
Panel A: dependent variable is Ln(real income per capita)				
Mine's wage bill and local purchases × distance < 100 km	0.174**	0.198**	0.186**	0.187**
	(0.078)	(0.085)	(0.086)	(0.087)
Observations	7738	6796	6668	6570
R-squared	0.521	0.484	0.478	0.479
Panel B: dependent variable is Ln(relative value of house rent)				
Mine's wage bill and local purchases × distance < 100 km	0.238***	0.240**	0.255***	0.256***
	(0.092)	(0.096)	(0.096)	(0.097)
Observations	7696	6829	6718	6613
R-squared	0.65	0.595	0.583	0.583
Public sector workers	yes	no	no	no
Mining workers	yes	yes	no	no
Live in district adjacent to mine pits	yes	yes	yes	no

Notes: Robust standard errors in parentheses. Standard errors are adjusted for clustering at primary sampling unit level. * denotes significant at 10%, ** significant at 5% and *** significant at 1%. All regressions include year and district fixed effects and the same control variables of the baseline regressions (see notes of Tables 3 and 7).

7 Additional Checks

In this section we present additional checks on the validity of the previous results. We are particularly interested on confounding factors that may violate the identification assumption, compositional effects driven by migration within the region and the sensitivity of the results to alternative measures of distance.

We focus on the baseline regression exploring the effect of the mine on real income. We also perform similar checks on the set of results on local crop prices and house rents and reach similar conclusions. These additional checks are not reported, but are available from the authors upon request.

7.1 Confounding Factors

A main concern is that, in parallel to the expansion of the mine, there are other phenomena happening in the region that affected differently areas close and far from the city. This would violate our identification assumption and imply that the estimated effect on real income could not be attributed to the mine expansion.

We are particularly concerned with changes on the survey sampling framework in 2001, heterogenous trends driven by different initial conditions or by proximity to a city.

7.1.1 Methodological Change

In 2001 some changes were introduced to the sampling framework of the household survey to include newly developed urban areas and improve the estimates of the poverty line (INEI, 2006). Our concern is that the change in methodology may have affected differently areas close and far to the mine and that our results are simply reflecting the change on the survey methodology and not the effects of the mine.

To address this issue, we include in our baseline regression a dummy equal to 1 for all years from 2001interacted with our measure of distance, to capture any shift in the average income level after 2001 for households closer to Cajamarca due to the re-sampling. Column (1) in Table 14 shows that our main result is not affected, suggesting that the income of adjacent households

due to the small size of the sample.

does not just suffer a jump in levels (as a change only due to the methodology would suggest) but follows the variation in mine's activity. If anything, accounting for this change in methodology increases the magnitude of our coefficient, suggesting that the bias was towards zero, rather than positive.

7.1.2 Heterogenous Trends

There are some systematic differences between areas close and far to Cajamarca. Areas closer to the city are relatively more urbanized, more densely populated and less agricultural. In the main specification, this is dealt with by using district fixed effects and controlling for household characteristics. However, these different initial conditions may also lead to different trends of income or prices which we may be mistakenly attributing to the mine expansion. For example, farmers' income may be diverging from non-farmers, or there may be an overall decline of rural areas or possibly densely populated areas are growing faster.

To address this concern we include a non-parametric trend interacted with dummies related to observable characteristics. In particular we use an indicator of urbanization (a dummy equal to 1 if the household is located in an urban area), agricultural activity (1 if the household reports any agricultural production), population density (1 if the population density of the district where the household lives is above the median).

Columns (2) to (4) in Table 14 shows the results of this robustness check using real income as dependant variable.[73] The estimates of the effect of the mine on real income are similar to those found in the baseline regression.

7.1.3 Proximity to a city

The previous checks suggest that our results are not driven by different trends based on some observable characteristics. There may be, however, unobservable shocks contemporaneous to the mine expansion that affect differently areas close or far from the city. For example, there may be a general process of divergence between cities and their rural hinterland due to agglomeration economies, different vulnerability to climatic shocks or targeting of public projects to main cities.

[73] The results are similar using other dependant variables studied in the main results.

Table 14: Effect of the mine on real income, controlling by methodological change and heterogenous trends

	Ln (real income per capita)			
	(1)	(2)	(3)	(4)
Mine's wage bill and local purchases × distance < 100 km	0.238**	0.168**	0.177**	0.184**
	(0.109)	(0.079)	(0.080)	(0.075)
Post 2001 × distance < 100 km	yes	no	no	no
Year dummies × urban	no	yes	yes	yes
Year dummies × farmer	no	no	yes	yes
Year dummies × high density	no	no	no	yes
Observations	7738	7738	7738	7738
R-squared	0.522	0.526	0.532	0.537

Notes: Robust standard errors in parentheses. Standard errors are adjusted for clustering at primary sampling unit level. * denotes significant at 10%, ** significant at 5% and *** significant at 1%. All regressions include year and district fixed effects and control variables. Post 2001 is a dummy equal to 1 if year is 2001 or later. Urban is a dummy equal to 1 if household resides in an urban area. Farmer is equal to 1 if household reports agricultural production. High density is equal to 1 if density of the district of residence is above the median in the sample.

We address this concern in two steps. First, we include as additional control to the baseline regression the measure of mine activity interacted with distance to other cities, to rule out that proximity to other city centers is driving the results. Second, we perform a falsification test replicating the estimates of the effect of the mine on real income but using as reference points other cities instead of Cajamarca. Finding a similar effect of the mine expansion on other cities would suggest that the observed effect on real income observed in areas closer to Cajamarca city is just reflecting a broader city-rural phenomenon and would raise concerns about the validity of the identification assumption

We select the other main cities around the North Highlands region: Chachapoyas, Chiclayo and Trujillo (see Figure 4 for a localization map of the cities). All these cities are department's capitals, as Cajamarca, and have a similar governmental status. Chachapoyas is located in the highlands and have a similar size as Cajamarca. In contrast, Chiclayo and Trujillo are much larger cities located on the coast. For each city, we calculate distance by road using the shortest path and construct a distance dummy variable equal to 1 if the distance is less than 100 km and 0 otherwise. The algorithm is the same used to calculate distance to Cajamarca city.

Table 15 presents the estimates of the baseline regression with real income as the dependant variable and including the mine wage bill and local purchases interacted with distance to other cities. In all cases, the mine expansion only affects areas closer to Cajamarca city, not to other urban centers.

Table 16 displays the results of the falsification test estimating the baseline regression with distance to the alternative city instead of Cajamarca and using two alternative samples. Panel A uses the same sample as in the baseline results: households in the North Highland region. Panel B includes households within 200 km of the cities, regardless of the geographical region.[74] In all cases the effect of mine wages and purchases becomes insignificant or even negative. The lack of effect on this falsification exercise weakens the case of some confounding factor driving the results.

[74] The results are similar using larger areas of influence, e.g. 400 km

Table 15: Effect of the mine on real income, controlling by proximity to other cities

	Ln(real income per capita)		
	(1)	(2)	(3)
Mine's wage bill and local purchases ×			
distance to Cajamarca	0.160*	0.160*	0.144*
< 100 km	(0.083)	(0.083)	(0.082)
distance to Chachapoyas	-0.088	-0.088	-0.103
< 100 km	(0.154)	(0.155)	(0.153)
distance to Chiclayo		0.014	0.001
< 100 km		(0.179)	(0.179)
distance to Trujillo			-0.042
< 100 km			(0.170)
Observations	7738	7738	7738
R-squared	0.520	0.520	0.520

Notes: Robust standard errors in parentheses. Standard errors are adjusted for clustering at primary sampling unit level. * denotes significant at 10%, ** significant at 5% and *** significant at 1%. All regressions include year and district fixed effects and control variables. Distance to all cities is calculated as the shortest path between the main town in the district where the household lives and the city.

Table 16: Falsification test using distance to other cities

	Ln(real income per capita)		
	(1)	(2)	(3)
Panel A: North Highlands sample			
Mine's wage bill and local purchases × distance < 100 km	-0.188	-0.066	-0.136
	(0.142)	(0.169)	(0.157)
Observations	7738	7738	7738
R-squared	0.521	0.521	0.521
Panel B: Households within 200 km			
Mine's wage bill and local purchases × distance < 100 km	-0.058	0.000	-0.077*
	(0.063)	(0.050)	(0.044)
Observations	6978	10801	12073
R-squared	0.552	0.536	0.604
City	Chachapoyas	Chiclayo	Trujillo

Notes: Robust standard errors in parentheses. Standard errors are adjusted for clustering at primary sampling unit level. * denotes significant at 10%, ** significant at 5% and *** significant at 1%. All regressions include year and district fixed effects and control variables.

7.2 Migration and Compositional Effects

The model predicts an increase on city size due to intra-regional migration. In the model, the increase on labor demand and wages in the city attracts agricultural workers. To explore this model prediction, we use data from two most recent population censuses (1993 and 2007) and calculate the intercensal population growth rate in Cajamarca city and surrounding areas, up to 400 km.[75]

Figure 13 plots the population growth rates for areas at different distance from the city. In the period between the two censuses, the region grew at an annual rate of 0.7%, below the national average of 2%. However, Cajamarca city experienced a significant growth, with population increasing at a rate of 3.4% per year and becoming the second fastest growing city in Peru (INEI, 2008).[76] In contrast, the surrounding areas grew at a much slower pace, below the regional average. Interestingly, the growth rate decreases with distance to the city.[77]

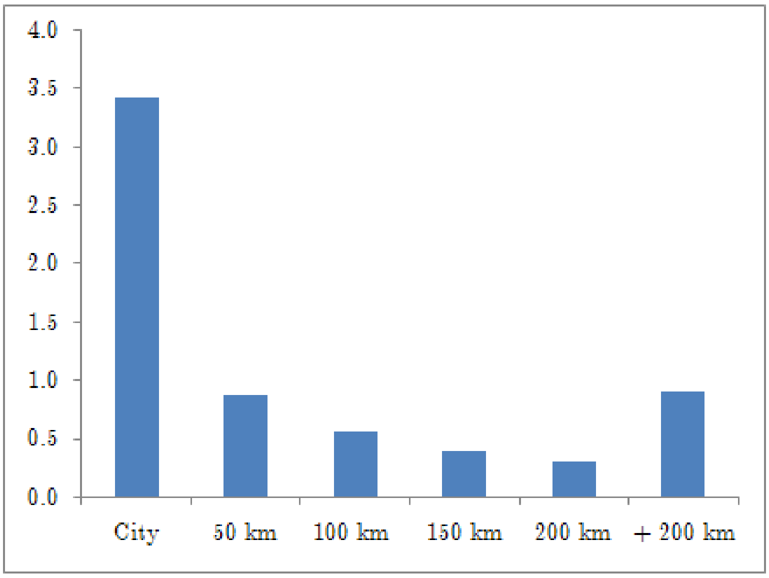

Figure 13: Intercensal annual population growth rate: Cajamarca city and surroundings

[75]This extension corresponds to the range of distances observed in the data.

[76]In the period between 1993 and 2007, Cajamarca city increased its population from 117,000 to almost 190,000.

[77]A possible explanation consistent with this observation is that migration costs are increasing with distance.

Taken together, this evidence is suggestive of migration from the rural hinterland to the city, as predicted by the model.[78] Empirically, a main concern with migration is that the observed increment on real income may be just reflecting compositional changes on the labor force. For example, if only the most productive agricultural workers migrate to the city, the increase on real income would be driven by higher productivity not by the demand shock from the mine.

We address this concern indirectly by evaluating whether the expansion of the mine has lead to changes on observable characteristics of the labor force in areas closer and farther from the city.[79] In particular we focus on different measures of human capital such as: years of education, indicators of having completed primary school or attended a private school (a proxy for quality of education).[80] We also consider demographics such as age, gender and number of household members to evaluate selective migration of young adult males. Finally, we explore characteristics of the agricultural unit such as proportion of production sold, number and concentration of crops, to account for changes on the type of farmers. In all cases, we estimate the baseline regression (20) with year and district fixed effect as the only control variables.

Table 17 shows the results. The parameter associated to the interaction term captures the changes on the measure of human capital, demographics or agricultural activity in areas closer to the city associated to the expansion of the mine. Note that all the measures of education (columns 1 to 3) worsen in areas closer to the city, while the while the rest of characteristics remain unchanged. If we take into account that educational attainment in areas far from the city was *lower* before the mine expansion, these results suggest a reduction on the gap between both areas due to the migration of relatively less educated workers from the rural hinterland to the city.[81] Taken together,

[78]Nonetheless, we cannot interpret it as evidence of the mine expansion *causing* migration because, among other reasons, we do not know whether migration occurred before or after the expansion of the mine.

[79]Ideally we would like to identify migrants in the sample and check whether the results are driven by this sub-population. Unfortunately, that information is not available.

[80]In the baseline regressions we control for education. However, this control may be insufficient to account for compositional changes in the presence of human capital spillovers or complementarities.

[81]In 1997 the average worker located more than 100 km from the city had 3 years of education while the average worker closer to the city had 3.6 years of education. In 2006, the years of education of both type of workers were 4.6 and 4.7, respectively

these results reduce concerns that the increase on real income is driven by migration of more productive workers to the city.

7.3 Alternative Measures of Distance

In this section we check that our results are not sensitive to our measure of distance. So far, we have used the shortest path by road and the average (i.e. 100 km) as the threshold. In column 1 of Table 18 we show that results hold when the threshold is defined by the median, i.e. 90 km.

Additionally, we obtain two alternative measures of distance: a topographic and a straight line. The topographic measure is calculated using the ArcGIS package by minimizing the sum of the normalized values of altitude and gradient, regardless of the existence of a road. It can be interpreted as a proxy for where a road may be located or alternative transportation routes in the absence of roads. The straight line measure is calculated as the Euclidian distance between the district capital town and the city of Cajamarca. In order to distinguish district closer and farther from the city, we use as a threshold the median value of the measure of distance. Columns 2 and 3 show that the effects are similar, irrespective of the measure of distance used to tell apart districts that are far and close to Cajamarca.

Finally, we explore in more detail the monotonic decline of the effect by distance, which is a crucial feature of our identification strategy. To do that we estimate the baseline regression including the interaction between the mine wage bill and local purchases and different functions of distance. Column 1 in Table 19 displays the results with the linear measure of distance as a benchmark. Columns 2 and 3 allow for non-linearities by including the logarithm and inverse of distance. In all cases, the results support the claim that the effect of the mine expansion on real income declines with distance to the city.

Table 17: Changes on characteristics of labor force and agricultural activity

	Human capital			Demographics		Agricultural activity		
	Years of education	Complete primary	Private school	Age	Is female	% prod. sold	Number of crops	Crop HH index
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Mine's wage bill and local purchases × distance < 100 km	-0.684**	-0.059*	-0.019***	-0.648	0.011	-0.094*	-0.153	0.015
	(0.322)	(0.031)	(0.007)	(0.835)	(0.017)	(0.050)	(0.287)	(0.036)
Observations	31255	31255	25473	34507	34159	5582	5582	5582
R-squared	0.106	0.064	0.177	0.023	0.011	0.309	0.311	0.228

Notes: Robust standard errors in parentheses. Standard errors are adjusted for clustering at primary sampling unit level. * denotes significant at 10%, ** significant at 5% and *** significant at 1%. All regressions include year and district fixed effects. Regressions in Columns 1 to 5 use the sample of individuals in working age, while columns 6 to 8 use the sample of households with some agricultural production. Complete primary is a dummy equal to 1 if individual completed primary school, private school is a dummy equal to 1 if individual assisted to private school. Is female is a dummy equal to 1 if individual is a female. Crop concentration is the Herfindahl-Hirschman concentration index calculated as the sum of squares of the contribution of a crop to total agricultural production.

Table 18: Alternative measures of distance

	Ln(real income per capita)		
	(1)	(2)	(3)
Mine's wage bill and local purchases × distance < median	0.137*	0.150**	0.134*
	(0.077)	(0.075)	(0.074)
Measure of distance	Shortest path by road	Topographic	Straight line
Median distance (km)	90	75	65
Observations	7738	7738	7738
R-squared	0.521	0.521	0.521

Notes: Robust standard errors in parentheses. Standard errors are adjusted for clustering at primary sampling unit level. * denotes significant at 10%, ** significant at 5% and *** significant at 1%. All regressions include year and district fixed effects and control variables. Topographic distance is calculated as the length of the shortest path that minimizes the normalized sum of altitude and gradient. Straight line distance is the Euclidean distance between two points.

Table 19: Exploring the decrease of the effect by distance

	Ln(real income per capita)		
	(1)	(2)	(3)
Mine's wage bill and local purchases x			
distance	-0.128**		
	(0.062)		
Ln(distance)		-0.034**	
		(0.016)	
distance^{-1}			0.002*
			(0.001)
Observations	7738	7738	7738
R-squared	0.521	0.521	0.521

Notes: Robust standard errors in parentheses. Standard errors are adjusted for clustering at primary sampling unit level. * denotes significant at 10%, ** significant at 5% and *** significant at 1%. All regressions include year and district fixed effects. Distance is equal to the length of the shortest path by road from the main town of the district where the household lives to Cajamarca city, expressed in hundreds kilometers.

8 Conclusion

This paper investigates the effect of a large mine on a regional economy using the analytical framework provided by a spatial general equilibrium model. We find robust evidence that the mine has generated positive income and net welfare gains for residents in the city and in the surrounding rural hinterland.

The main contribution of the paper is to improve the understanding of the mechanisms through which natural resource extraction can foster local development. In particular it shows that, in the presence of backward linkages, the expansion of extractive industries can generate a positive demand shock and increase the real return to local factors of production, such as land and labor. In turn, this translates into better living conditions for local residents.

A main limitation of the paper is that we only observe events occurring over the span of a decade during the mine operation. This means that we are unable to explore whether the welfare gains are a short-term effect or part of sustainable development that would persist after the mine closure. For the same reason, we can say little about relevant long-run phenomena such as specialization, technological progress, or city formation (such as San Francisco after the Gold Rush). Though beyond the scope of this paper, these phenomena warrants further research.

In the case we study, the positive effects come from a market channel rather than from the revenue windfall to local governments. This suggests that, in a context of weak governments, policies that promote local procurement and employment could be more beneficial to local residents than increased public spending, at least in the short run.

This policy implication, however, depends of the pre-existence of good and labor markets o able to supply local inputs to the mine, and the extent of economic integration of the region where the mine is located. In the absence of inter-regional trade, the benefits would disappear since the increase on good prices could offset the gains on nominal income. Similarly, lack of intra-regional trade or labor mobility would restrict the benefit to the place where the mine suppliers are located and might lead to negative re-distributional effects.

The availability of natural resources in the developing world is often seen as a hindrance to economic development. In most cases, institutional failure (e.g. conflict, mismanagement or corruption) is at the heart of this inability to transform natural wealth into better standards of living. However, as this paper suggests, in the presence of strong enough backward linkages natural resources can be more a blessing than a curse.

References

Banco Central de Reserva del Perú, "Cuadros Anuales Históricos," 2009.

Benjamin, Nancy C., Shantayanan Devarajan, and Robert J. Weiner, "The 'Dutch' disease in a developing country : Oil reserves in Cameroon," *Journal of Development Economics*, January 1989, *30* (1), 71–92.

Black, Dan, Terra McKinnish, and Seth Sanders, "The Economic Impact of the Coal Boom and Bust," *Economic Journal*, April 2005, *115* (503), 449–476.

Black, Sandra E., "Do Better Schools Matter? Parental Valuation of Elementary Education," *Quarterly Journal of Economics*, 1999, *114* (2), 577–599.

Carrington, William J, "The Alaskan Labor Market during the Pipeline Era," *Journal of Political Economy*, February 1996, *104* (1), 186–218.

Caselli, Francesco, "Power Struggles and the Natural Resource Curse," 2006.

____ **and Guy Michaels**, "Do Oil Windfalls Improve Living Standards? Evidence from Brazil," June 2009.

Corden, W. Max and J. Peter Neary, "Booming Sector and De-Industrialisation in a Small Open Economy," *The Economic Journal*, 1982, *92* (368), 825–848.

Deaton, Angus, *The Analysis of Household Surveys: A Microeconometric Approach to Development Policy*, John Hopkins University Press for The World Bank, 1997.

Evans, William N. and Julie H. Topoleski, "The Social and Economic Impact of Native American Casinos," NBER Working Papers 9198, National Bureau of Economic Research September 2002.

Fujita, Masahisa and Paul Krugman, "When is the economy monocentric?: von Thünen and Chamberlin unified," *Regional Science and Urban Economics*, August 1995, *25* (4), 505–528.

——— , ——— , **and Anthony J. Venables**, *The Spatial Economy: Cities, Regions, and International Trade*, The MIT Press, 1999.

Greenstone, Michael and Enrico Moretti, "Bidding for Industrial Plants: Does Winning a 'Million Dollar Plant' Increase Welfare?," NBER Working Papers 9844, National Bureau of Economic Research, Inc July 2003.

——— **and Justin Gallagher**, "Does Hazardous Waste Matter? Evidence from the Housing Market and the Superfund Program," *Quarterly Journal of Economics*, 2008, *123* (3), 951–1003.

Hirschman, Albert O., *The Strategy of Economic Development*, Yale University Press, 1958.

Hooker, Mark A and Michael M Knetter, "Measuring the Economic Effects of Military Base Closures," *Economic Inquiry*, October 2001, *39* (4), 583–98.

Humphreys, David, "Emerging Players in Global Mining," Extractive Industries and Development Series 5, The World Bank June 2009.

INEI, "Condiciones de Vida en el Perú: Evolución 1997-2004," 2006.

——— , "Perú: Perfil de la Pobreza según departamentos 2004-2006," 2007.

——— , "Perú: Compendio Estadístico," 2008.

Jenkins, Beth, Anna Akhalkatsi, Brad Roberts, and Amanda Gardiner, "Business Linkages: Lessons, Opportunities and Challenges," Technical Report, IFC, International Business Leaders Forum, and the Fellows of Harvard College.

Koenker, Roger and Kevin F. Hallock, "Quantile Regression," *Journal of Economic Perspectives*, Fall 2001, *15* (4), 143–156.

Kuramoto, Juana R., "Las aglomeraciones productivas alrededor de la minera: el caso de la minera Yanacocha S.A," Documentos de Trabajo 27, Grupo de Análisis para el Desarrollo (GRADE) July 1999.

Lederman, Daniel and William F. Maloney, "Trade structure and growth," Policy Research Working Paper Series 3025, The World Bank April 2003.

Magee, L., A. L. Robb, and J. B. Burbidge, "On the use of sampling weights when estimating regression models with survey data," *Journal of Econometrics*, June 1998, *84* (2), 251–271.

Michaels, Guy, "The Long-Term Consequences of Regional Specialization," CEPR Discussion Papers 6028, CEPR Discussion Papers December 2006.

Minera Yanacocha, "Balance Social," 2002.

___ , "Balance Social y Ambiental," 2006.

Ministerio de Energía y Minas, "Anuario Minero," 2006.

Pascó-Font, Alberto, Alejandro Diez Hurtado, Gerardo Damonte, Ricardo Fort, and Guillermo Salas, "Chapter 4. Peru: Learning by Doing," in Gary McMahon and Felix Remy, eds., *Large Mines and the Community: Socioeconomic and Environmental Effects in Latin America, Canada and Spain*, The World Bank, 2001.

Roback, Jennifer, "Wages, Rents, and the Quality of Life," *Journal of Political Economy*, December 1982, *90* (6), 1257–78.

Sachs, Jeffrey D. and Andrew M. Warner, "Natural Resource Abundance and Economic Growth," NBER Working Papers 5398, National Bureau of Economic Research, Inc December 1995.

___ **and** ___ , "The big push, natural resource booms and growth," *Journal of Development Economics*, June 1999, *59* (1), 43–76.

___ **and** ___ , "The curse of natural resources," *European Economic Review*, May 2001, *45* (4-6), 827–838.

Sala-i-Martin, Xavier, Gernot Doppelhofer, and Ronald I. Miller, "Determinants of Long-Term Growth: A Bayesian Averaging of Classical Estimates (BACE) Approach," *American Economic Review*, September 2004, *94* (4), 813–835.

Torvik, Ragnar, "Natural resources, rent seeking and welfare," *Journal of Development Economics*, April 2002, *67* (2), 455–470.

World Bank, "A Sourcebook for Poverty Reduction Strategies: Mining," 2002.

A Proofs

A.1 Existence of equilibrium

Assumption 5 $N > \frac{1-\alpha}{1-\alpha-\mu}\theta S$

This assumption means that the share of employment from the mine is small enough relative to the total population in the region. To check how reasonable this assumption is, note that we can re-write it as $\mu < (1-\alpha)\left(1-\frac{\theta S}{N}\right)$, where α is the consumer's budget share of locally produced food, μ is the budget share of local manufactures and $\frac{\theta S}{N}$ is the relative size of the mine's workforce.

We calculate the budget share on food using the ENAHO survey and use it as a proxy for α.The estimated budget share is 0.6, however note that this measure may overestimate the true value of α since some of the food is not locally produced. Regarding $\frac{\theta S}{N}$, recall from Section 3 that Yanacocha's direct and indirect employment in the analyzed period was between 12% and 20% of the active population of the city of Cajamarca. This proportion would be significantly lower if we include working population in the rest of the region. These figures imply that for Assumption 1 to be violated μ should be, at least, greater than 0.32 or 0.35. In turn this implies a very small budget share of imported goods $(1-\alpha-\mu)$, with values at most between 0.05-0.08.[82]

Proposition 6 *Under Assumption 5, there is a unique and positive pair* $(f^*; w^*)$ *that solves* (AA) *and* (BB).

Proof. To see this, first note that we can re-write equilibrium condition (BB) as a function of $L_e(w)$ and plug it in equilibrium condition (AA). We then obtain an expression of f only that can be expressed as

$$H(f) = N - 2c_a f - \frac{1-\alpha}{1-\alpha-\mu}\theta S - A(f)$$

,where $A(f) = (1 + \frac{\mu}{1-\alpha-\mu}\frac{1}{\varepsilon})\frac{\varepsilon}{1-\varepsilon}\{2c_a[\frac{(1-\alpha)(1-e^{-\tau f})}{\alpha e^{-2(1-\alpha)\tau f}} + f] - N\}$, $A'(f) > 0$ and $H'(f) < 0$. To find the equilibrium, we look for a value f^* such that $H(f^*) = 0$.

[82]Note that the overestimation of α and $\frac{\theta S}{N}$ mean that the threshold that μ should exceed to violate assumption 1 are understated. Hence, violation of assumption 1 requires an even higher μ, and smaller budget share of imported goods.

First, recall that the values that f can adopt in equation (AA) have an upper bound $\overline{f} = \frac{1}{2c_a}\left(N - \frac{1-\alpha}{1-\alpha-\mu}\theta S\right)$ with $0 < \overline{f} < \dfrac{N}{2c_a}$ under Assumption 5. Note that $H(\overline{f}) = -A(\overline{f})$, which is negative only if $2c_a[\frac{(1-\alpha)(1-e^{-\tau\overline{f}})}{\alpha e^{-2(1-\alpha)\tau\overline{f}}} + \overline{f}] - N > 0$. From definition of $\overline{f}$, (11) and (16), this condition can be re-written as $\frac{1-\alpha}{1-\alpha-\mu}\theta S + \dfrac{\pi_e}{w} > \frac{1-\alpha}{1-\alpha-\mu}\dfrac{\theta S}{2c_a}$, which is satisfied under Assumption 5, hence $H(\overline{f}) < 0$.

Second, recall that the values of f in equation (BB) have a lower bound $\underline{f} \equiv \lim\limits_{w\to\infty} f = g^{-1}(N)$, where $g^{-1}(\cdot)$ is the inverse function of $g(f) \equiv 2c_a[\frac{(1-\alpha)(1-e^{-\tau f})}{\alpha e^{-2(1-\alpha)\tau f}} + f]$. from this definition follows that $A(\underline{f}) = 0$. Combining this result with the previous observations that $A(\overline{f}) > 0$ and $A'(f) > 0$, it follows that $0 < \underline{f} < \overline{f} < \dfrac{N}{2c_a}$ and therefore $H(\underline{f}) = N - 2c_a\underline{f} - \frac{1-\alpha}{1-\alpha-\mu}\theta S > 0$.

Finally, since $H(f)$ monotonically decreases in f, $H(\underline{f}) > 0$ and $H(\overline{f}) < 0$ imply that there is a unique positive value $f^* \in (\underline{f}, \overline{f})$, such that $H(f^*) = 0$. From inspection of either (11) or (16), values of $f^* \in (\underline{f}, \overline{f})$ imply a value of $L_e(w) > 0$ and thus that the equilibrium wage, namely w^*, is also unique and positive. ■

A.2 Effect on number of manufacturing firms

Proposition 7 *the number of manufacturing firms, v increases with the expansion of the mine, $\frac{dv}{dS} > 0$, if and only if $\theta > -\dfrac{1}{\varepsilon}\dfrac{dL_e}{dw}\dfrac{dw}{dS}$.*

Proof. Taking total derivatives to (AA) we obtain that:

$$-2c_a\frac{df}{dS} - \frac{(1-\alpha-\mu)\,\varepsilon + \mu}{(1-\alpha-\mu)\,\varepsilon}\frac{dL_e}{dw}\frac{dw}{dS} = \frac{1-\alpha}{1-\alpha-\mu}\theta. \tag{22}$$

Similarly taking total derivatives to (17) we obtain

$$\frac{dv}{dS} = -2c_a\frac{df}{dS} - \frac{dL_e}{dw}\frac{dw}{dS} - \theta \tag{23}$$

and thus $\frac{dv}{dS} > 0$ if and only if:

$$-2c_a\frac{df}{dS} - \frac{dL_e}{dw}\frac{dw}{dS} > \theta. \tag{24}$$

Using (22) and (9) we can re-write condition (24) as $\theta > -\dfrac{1}{\varepsilon}\dfrac{dL_e}{dw}\dfrac{dw}{dS}$. ■

A.3 Effect on real wage

Proposition 8 *if the number of manufacturing firms increase with expansion of the mine, then* $\frac{d\omega}{dS} > 0$*. Otherwise, the effect is more likely to be positive for low values of* μ*,* θ *or high values of* ε*.*

Proof. Taking total derivatives to (18) and using (13) we obtain:

$$\frac{d\omega}{dS} = \frac{\mu}{\sigma-1}(\frac{w}{G})^{\mu}v^{-1}\frac{dv}{dS} + 2(1-\alpha)\alpha\tau(\frac{w}{p_a})^{\alpha}(-\frac{df}{dS}) + \\ (1-\alpha-\mu)(\frac{w}{p_a})^{1-\alpha-\mu}w^{-1}\frac{dw}{dS}, \tag{25}$$

which is positive if $\frac{dv}{dS} > 0$.

To study the conditions for $\frac{d\omega}{dS} > 0$ when $\frac{dv}{dS} < 0$, we first re-write (25) replacing $\frac{dv}{dS}$ using expression (23):

$$\frac{d\omega}{dS} = C - \frac{dL_e}{dw}\frac{dw}{dS} - \theta,$$

where

$$C \equiv -\{2(1-\alpha)\alpha\tau(\frac{w}{p_a})^{\alpha}\frac{\sigma-1}{\mu}(\frac{w}{G})^{-\mu}v + c_a\}\frac{df}{dS} + \\ (1-\alpha-\mu)(\frac{w}{p_a})^{1-\alpha-\mu}w^{-1}\frac{\sigma-1}{\mu}(\frac{w}{G})^{-\mu}v\frac{dw}{dS}.$$

Hence the condition for $\frac{d\omega}{dS} > 0$ when $\frac{dv}{dS} < 0$ is:

$$C - \frac{dL_e}{dw}\frac{dw}{dS} > \theta.$$

Note, from Proposition 7, that $\frac{dv}{dS} < 0$ implies $-\frac{dL_e}{dw}\frac{dw}{dS} > \varepsilon\theta$. That means that a sufficient condition for $\frac{d\omega}{dS} > 0$ when $\frac{dv}{dS} < 0$ is:

$$C > \theta(1-\varepsilon).$$

This condition is more likely to be satisfied for high values of ε or low values of μ and θ. ■

A.4 Proof of Propositions 1 and 2

Using the definition of real income from (19) and re-arranging, we can re-write the food market equilibrium condition (11) as:

$$\frac{p_a}{P} = \left[2\left(1-\alpha\right)\int_0^f e^{-\tau r}dr\right]^{-1} \alpha L y,$$

where y is the real income in the city. Note that $\frac{dy}{dS} > 0$ implies $\frac{dp_a/P}{dS} > 0$, because $\frac{df}{dS} < 0$ and $\frac{dL}{dS} > 0$. However if $\frac{dy}{dS} > 0$ the sign of $\frac{dp_a/P}{dS}$ is ambiguous.

Recall that real income in the rural hinterland is:

$$y(r) = \frac{1}{k}\frac{p_a}{P}e^{-2(1-\alpha)\tau|r|}.$$

It is immediate to see that $\frac{dy(r)}{dS} > 0$ and $\frac{d^2y(r)}{dSdr} < 0$.

A.5 Proof of Proposition 3

Using (2) and (5), we can write the relative prices of food and the import good as:

$$\begin{aligned} \frac{p_a(r)}{P(r)} &= \frac{p_a}{P}e^{-2(1-\alpha)\tau|r|} \\ \frac{p_m(r)}{P(r)} &= \frac{p_m}{P}e^{2\alpha\tau|r|}. \end{aligned}$$

By proposition 1, we know that if $\frac{dy}{dS} > 0$, then $\frac{dp_a/P}{dS} > 0$ which also implies that $\frac{dp_a(r)/P(r)}{dS} > 0$ Since $\frac{dP}{dS} \lesseqgtr 0$ and $\frac{dp_m}{dS} = 0$, it follows that $\frac{dp_m(r)/P(r)}{dS} \lesseqgtr 0$.