Natural Resource Booms and Inequality: Theory and Evidence

Goderis, Benedikt and Malone, Samuel W.

University of Oxford

28 July 2009

Natural Resource Booms and Inequality: Theory and Evidence*

Benedikt Goderis †
University of Oxford

Samuel W. Malone ‡
University of the Andes

July, 2009

Abstract

Surprisingly little is known about the impact of natural resource booms on income inequality in resource rich countries (Ross, 2007). This paper develops a theory, in the context of a two sector growth model in which learning-by-doing drives growth, to explain the time path of inequality following a resource boom. Under the condition that the nontraded sector uses unskilled labor more intensively than the traded sector, we find that income inequality will fall in the short run immediately after a boom, and will then increase steadily over time as the economy grows, until the initial impact of the boom on inequality disappears. Using dynamic panel data estimation for 90 countries between 1965 and 1999, and exploiting variation in world commodity prices to identify resource booms, we find evidence in support of the theory, especially for oil and mineral booms. We also find that uncertainty about future commodity export prices significantly increases long-run inequality.

Keywords: Resource Booms; Inequality; Dutch Disease

JEL Classification: O13, O15, F11, Q33

*We would like to thank Tony Atkinson, Paul Collier, Markus Eberhardt, Sandeep Kapur, Peter Neary, Ron Smith, Ragnar Torvik, Rick van der Ploeg, David Vines, and two anonymous referees for many helpful comments.

†Department of Economics, University of Oxford, Manor Road, Oxford OX1 3UQ, UK. Tel.: +44-(0)1865-281440, Fax: +44-(0)1865-281447, Email: Benedikt.Goderis@economics.ox.ac.uk.

‡School of Management, University of the Andes, Bogota, Colombia, Calle 21 No. 1-20-piso 9. Email: samuel.malone@gmail.com.

1 Introduction

Large windfalls from periodical commodity booms pose important questions for policy makers in natural resource-rich economies. The question that has received most attention is how the windfalls can be used to promote *economic growth*. However, surprisingly little is known about the impact of booms on *income inequality* (Ross, 2007). The persistence of high poverty rates in many resource-rich economies makes the question of how booms affect the distribution of income between the rich and the poor particularly relevant.

A useful starting point for our analysis is the theory of Dutch Disease. The resource revenues lead to an appreciation of the real exchange rate, which harms the competitiveness of the non-resource exports sector and hampers economic growth if there are positive externalities to production in this sector (Corden and Neary, 1982; Van Wijnbergen, 1984; Torvik, 2001). We develop a theory, in the context of a two sector growth model in which learning-by-doing drives growth, to explain the time path of inequality following a natural resource boom (defined as either a discovery or an exogenous world price increase). Under the plausible conditions of low elasticity of substitution between nontraded and traded goods in consumption, a nontraded sector that is relatively intensive in its use of unskilled labor, and balanced growth, we find that income inequality will fall in the short run immediately after a boom, and will then increase steadily over time as the economy grows, until the initial impact of the boom on inequality disappears.

We then test the predictions of the model empirically, using a dynamic panel data estimator for 90 countries between 1965 and 1999, and exploiting exogenous variation in world commodity prices to identify natural resource booms. Our results support the theory. Resource booms, especially oil and mineral booms, lower inequality in the year of the boom. This effect then gradually diminishes over time until inequality returns to its pre-boom level in the long run. In addition to the estimated effects of resource booms, we also find that uncertainty about future commodity export prices significantly increases long-run inequality.

This paper is loosely related to the literature on natural resource endowments and inequality. Leamer et al. (1999) argue that, since resource exploitation does not require much human capital, the labor force in resource rich economies is unprepared for the emergence of human-capital-intensive manufacturing. As a result, these economies may experience higher income inequality for longer periods than resource-scarce economies. Sokoloff and Engerman (2000) instead point at how resource endowments affect inequality through the evolution of institutions. In colonies where economies of scale led to unequal land ownership, the inequality was sustained by political institutions that favored the rich and excluded the poor. In other colonies, however, the absence of economies of scale led to a more equal land distribution and more egalitarian institutions. Gylfason and Zoega (2003) argue that resource dependence leads to both lower growth and increased inequality, and could therefore explain the inverse relationship between growth and inequality in cross-country data. Finally, this paper is related to Ross (2007), who discusses mechanisms through which mineral wealth could affect inequality.

The rest of this paper is organized as follows. Section two presents the model and analyzes the effects of a resource boom on inequality in the short and long run. Section three describes the empirical analysis. Section four discusses the estimation results. Section five concludes.

2 The Model

Our model of the dependent economy has a nontraded sector, which we call the N sector, and a non-resource traded sector, which we call the T sector. The resource sector of the model is represented as an exogenous gift of resource income R.[1]

We normalize the population to be equal to 1. The population consists of L unskilled workers and S skilled workers, with $L + S = 1$. Both skilled and unskilled labor, the only factors of production, are required to produce N and T goods, and the factor stocks are fixed at L and S,

[1] In specifying the resource income in this manner, we eliminate the possibility of the resource movement effect (Cordon and Neary, 1982) to play a role in the adjustment of the economy to shocks, and focus exclusively on the spending effect, which has been acknowledged widely as the main force driving adjustment in most resource rich countries. It is worth noting, however, that inclusion of the former effect in the model, in our case by allowing skilled but not unskilled workers to be employed competitively in the resource sector, would likely dampen somewhat our result on the short run fall in inequality.

as noted above. Letting L_N and L_T represent the unskilled labor force in the N and T sectors, respectively, the market clearing condition for unskilled labor can be written as $L_N + L_T = L$. Similarly, the market clearing condition for skilled labor can be written as $S_N + S_T = S$, where S_N and S_T represent the skilled labor force in the respective sectors.[2]

To close the short run specification of the model, in which productivity in each sector is constant, we must specify the production functions, profit maximization conditions, preferences, and demand functions. For simplicity, we will assume Cobb Douglas production functions in each sector, with a unit elasticity of substitution between factors and constant returns to scale. Denoting output and productivity in the sectors by X_j and A_j, respectively, we have

$$X_N = A_N S_N^{\theta_{SN}} L_N^{\theta_{LN}} \quad \text{and} \quad X_T = A_T S_T^{\theta_{ST}} L_T^{\theta_{LT}}. \tag{1}$$

Constant returns to scale implies that $\theta_{Sj} + \theta_{Lj} = 1$, $j = N, T$. We make the assumption henceforth that the nontraded sector is relatively intensive in its use of unskilled labor: $\theta_{LN} > \theta_{LT}$. Additionally, we define λ_{ij} as the proportion of the supply of factor i used by sector j in equilibrium. Market clearing requires that $\lambda_{iN} + \lambda_{iT} = 1$, $i = L, S$.

Both unskilled and skilled labor must earn their marginal products, and assuming perfect factor mobility, the marginal product of each factor must be equal across sectors. This gives the wage w of the unskilled worker and the wage v of the skilled worker, respectively, as

$$w = p_N X_N'(L_N) = X_T'(L_T) \quad \text{and} \quad v = p_N X_N'(S_N) = X_T'(S_T), \tag{2}$$

where p_N is the relative price of nontraded goods in terms of traded goods.

All agents are assumed to have identical preferences, and maximize a CES aggregator of N and

[2] We focus on skilled and unskilled labor as factors of production in order to analyze the effects of resource booms on the *personal* income distribution. Additionally, our model of long run growth driven by learning-by-doing is framed more naturally in a setting where both types of workers are capable of LBD in both sectors. Our model is formally equivalent to a version formulated using capital and labor, however, and all our core results carry through so long as we maintain the assumption that the traded sector uses skilled labor (capital) more intensively than the nontraded sector, or equivalently that the nontraded sector uses unskilled labor (labor) more intensively than the traded sector.

T good consumption, which is given by

$$U = \left[(1-\gamma)^{1/\sigma} C_T^{\frac{\sigma-1}{\sigma}} + (\gamma)^{1/\sigma} C_N^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}} \tag{3}$$

Here σ is the elasticity of substitution between traded and nontraded goods. Since agents have identical preferences, we can treat aggregate consumption of the two goods as being determined by the decision of a representative agent who chooses C_N and C_T to maximize U subject to the budget constraint of the entire economy,

$$p_N C_N + C_T = Y. \tag{4}$$

Here $Y = p_N X_N + X_T + A_T R$ is aggregate income. A boom will be considered as an increase in R. It is first important to note that resource income will be measured in the productivity units A_T of the traded sector, as in Torvik (2001). This detail is irrelevant in the short run, when productivity in both sectors is constant, but will be highly relevant in the transition to the long run, when productivity in both sectors is growing. Second, the assumption that skilled and unskilled workers have identical tastes rules out the possibility for demand composition effects to play any role in the economy's response to booms. As in Obstfeld and Rogoff (1996), the overall price index in the economy, P, for the consumption basket that is the solution to the consumer's problem is given by

$$P = \left[1 - \gamma + \gamma p_N^{1-\sigma} \right]^{\frac{1}{1-\sigma}} \tag{5}$$

The demand functions for traded and nontraded goods respectively are given by

$$C_T = (1-\gamma) P^{\sigma-1} Y \quad \text{and} \quad C_N = \gamma P^{\sigma-1} p_N^{-\sigma} Y. \tag{6}$$

The model is closed by the requirement that the market for nontraded goods must clear: $X_N = C_N$. By Walras' Law, we can omit the market equilibrium for traded goods.

2.1 Measuring Inequality

There are two main sources of changes in income inequality during resource booms that concern us: the unequal distribution of resource income, and the shift of the factors of production to the nontraded sector, which uses unskilled labor intensively relative to the traded sector, due to the spending effect related to the resource income. We will first define a natural measure of non-resource income inequality based on the factor distribution of income, and relate this measure to the Gini coefficient of non-resource income inequality in our economy. Then, we will discuss how our results on changes in non-resource income inequality must be modified to take into account inequality in the receipt of resource income between unskilled and skilled workers.

The natural measure of non-resource income inequality we will focus on is the ratio of the total wage income earned by skilled workers to the total wage income earned by unskilled workers:

$$I = \frac{vS}{wL}. \tag{7}$$

This structural measure of non-resource income inequality captures the value share of skilled labor in producing output as a multiple of the value share of unskilled labor in producing output, with an emphasis that is consistent with the empirical findings of Daudey and García-Peñalosa (2007) that relative factor returns are a key driver of the inequality of the personal income distribution. Lemma 1 relates this measure of inequality to the Gini coefficient of non-resource income inequality:

Lemma 1 *Let the Gini coefficient of non-resource income inequality be denoted by G. If $w < v$, then $G = L - 1/(1 + I)$, and if $w \geq v$, then $G = S - I/(1 + I)$.*

Proof. The result follows from the definition of the Gini coefficient after routine calculations. ■

We will assume throughout the rest of the analysis that $w < v$, as consistent with the fact that skilled labor is relatively scarce and commands a higher wage. In this context, we see that G is a nonlinear, increasing function of I for a given unskilled labor supply L. Thus, the above lemma shows that our results on the sign of changes in I in the model can be used immediately to sign changes in the non-resource income Gini coefficient G, in the short run and the long run.

Since we are interested in the effects of resource booms on total income inequality, we next address the distribution of *resource* income among skilled and unskilled workers. Suppose that each unskilled worker earns total income $w + \alpha A_T R$ and each skilled worker earns total income $v + \beta A_T R$, where $A_T R$ is the value of the resource income measured in productivity units of the traded sector. Let the Gini coefficient of total income inequality be denoted by G^T and the Gini coefficient of resource income inequality be denoted by G^R. Let $y^{NR} \equiv wL + vS = p_N X_N + X_T$ denote the total value of non-resource income, let $y^R \equiv A_T R$ denote the total value of resource income, and let $y = y^{NR} + y^R$ denote the total income of the economy. Then we have the following:

Lemma 2 *For any $\alpha > 0$ and $\beta > 0$ such that $\alpha L + \beta S = 1$, the Gini coefficient G^T of total income inequality in the economy can be represented as follows, for the three possible cases:*

1. *If $\alpha \leq \beta$, then $G^T = (y^{NR}/y)G + (y^R/y)G^R$.*

2. *If $\alpha > \beta$ and $A_T R < (v - w)/(\alpha - \beta)$, then $G^T = (y^{NR}/y)G + (y^R/y)G^R - \frac{2R(\alpha-\beta)LS}{y}$*

3. *If $\alpha > \beta$ and $A_T R \geq (v - w)/(\alpha - \beta)$, then $G^T = (y^{NR}/y)G + (y^R/y)G^R - \frac{2(v-w)LS}{y}$.*

Here G is the non-resource Gini coefficient given in Lemma 1 for the case $w < v$, and the Gini of resource income inequality is given by $G^R = 1 - (\beta S^2 + 2\min[\alpha, \beta]LS + \alpha L^2)$.

Proof. The result follows from using the definitions of the total worker incomes, and the definition of the Gini coefficients constructed for the total, non-resource, and resource incomes of the individual workers, with attention paid to how the assumptions that define each case induce ordering of the total, non-resource, and resource incomes, respectively. ■

2.2 The Short Run Effect of a Resource Boom on Income Inequality

It is a natural assumption that the movement of factors between sectors in response to the return differentials generated by a resource boom, or increase in R, will occur on a much faster timescale than the productivity growth associated with changes in A_N and A_T over time. Before discussing the dynamics of productivity growth, therefore, it is worth stating the response of G and G^T to a resource boom when productivity levels in the two sectors are constant. We focus on the case of perfect factor mobility for both skilled and unskilled labor. This case is discussed in Corden and Neary (1982), section III, in a model with capital and labor. The results carry over directly to our measure of I in a model with skilled and unskilled labor. The net result on our measure of inequality G therefore follows directly from Corden and Neary (1982) and Lemma 1 above:

Proposition 3 *We will have $dG/dR < 0$ if and only if $\theta_{LN} > \theta_{LT}$.*

When spending effects are the primary driver of short run adjustment to resource booms, non-resource income inequality is likely to fall because in most countries, the nontraded sector, usually identified with services, real estate, and sometimes agriculture, is likely to use unskilled labor more intensively than the traded sector, which is usually identified with manufacturing.[3]

Unless the resource sector income constitutes a significant proportion of total economy activity, the behavior of G in response to resource booms will be the predominant driver of changes in G^T, the Gini coefficient of total income inequality, and we can expect G^T to fall in the short run in response to resource booms. In the general case, however, the fall in G will not be sufficient to guarantee the fall in G^T. Lemma 2 allows us to decompose the change in the Gini coefficient of total income inequality in response to a resource boom into three effects, due respectively to a change in the weighting of G and G^R, the change in G itself, and a change in the corrective term in cases 2 and 3. The following proposition summarizes a set of sufficient conditions that guarantee a fall in the Gini of total income inequality in response to a boom:

Proposition 4 *In cases 1 and 2 of Lemma 2, $\theta_{LN} > \theta_{LT}$ and $G^R < G$ are sufficient conditions to ensure that $dG^T/dR < 0$, so that resource booms induce a fall in the Gini coefficient of total income inequality. In case 3 of Lemma 2, $\theta_{LN} > \theta_{LT}$ and $G^R < G'$ are sufficient conditions to ensure that $dG^T/dR < 0$, where $G' \equiv G - 2L(I - S/L)/(1 + I)$.*

Proof. The result of case 1 follows directly by computing $dG^T/dR = (d(y^R/y)/dR)(G^R - G) + (1 - y^R/y)(dG/dR)$ and applying proposition 3. The result of case 2 follows from the case 1 result, plus the fact that $d(2(R/y)(\alpha - \beta)LS)/dR > 0$. The case 3 result follows by rearranging the corrective term in order to re-write the formula for G^T as $G^T = (y^{NR}/y)G' + (y^R/y)G^R$, showing that $dG'/dI > 0$, and applying the result from case 1. ■

[3] Although it does not bear on the empirics, it can be shown that the above result generalizes to the case of factor specificity, and that only the magnitude of the change in inequality, rather than the sign, is affected by relative factor mobility. A proof of these claims is available from the authors upon request.

In case 1, we see that our theoretical prediction of a short run fall in G carries over to a short run fall in total inequality G^T as well, and in fact, the fall in G^T will be *greater* than the fall in G. Moreover, the conditions given in Proposition 4 are sufficient, but not necessary, for a fall in the total income Gini; G^T will still fall in scenarios where the sufficient conditions stated above fail to be satisfied at the margin. In fact, even when skilled workers receive more rents per capita than unskilled workers ($\alpha > \beta$), and $G^R > G$, the fact that $dG/dR < 0$ for $\theta_{LN} > \theta_{LT}$ implies that we must have G^R greater than some endogenous threshold, which is decreasing in R, in order to have $dG^T/dR > 0$. The significant proportion of resource rents in total economic activity, and the high inequality of the distribution of resource rents in favor of skilled workers, required to produce a rise in the total income Gini seem unlikely to be satisfied by most resource rich countries in practice.

2.3 The Long Run Effect of a Resource Boom on Income Inequality

To analyze the transition of the economy to the long run, we relax the assumption of constant TFP levels in the N and T sectors and specify dynamic equations governing productivity growth.

The rate of productivity growth in each sector will be endogenously determined, and we will consider the effects of learning-by-doing (LBD) in both sectors with the possibility of knowledge spillovers. The model in this section is a generalization of Torvik (2001), who considers productivity growth in a specific factors model without an explicit consideration of the role of skilled versus unskilled labor. Including both factors makes it possible to study the level of inequality in the balanced growth steady state and in response to resource booms.

The evolution of productivity in each sector is given by the following pair of differential equations, which generalize the specification of Torvik (2001) to include the possibility of productivity growth driven by LBD on the part of both skilled and unskilled workers in each sector:

$$\frac{\dot{A}_N}{A_N} = u_L\lambda_{LN} + u_S\lambda_{SN} + \delta_T(v_L\lambda_{LT} + v_S\lambda_{ST}) \tag{8}$$

$$\frac{\dot{A}_T}{A_T} = \delta_N(u_L\lambda_{LN} + u_S\lambda_{SN}) + v_L\lambda_{LT} + v_S\lambda_{ST} \tag{9}$$

Productivity growth is driven by LBD in both sectors, and the amount of LBD in a given sector is determined by the level of employment of both types of workers in that sector. The strength of the effect of unskilled workers on productivity growth in a given sector, however, may differ from the strength of the effect generated by skilled workers on productivity growth in the sector. The direct effect of unskilled worker employment on productivity growth in the N sector is measured by the parameter u_L, and the direct effect of skilled worker employment on N sector productivity growth is measured by the parameter u_S. Similarly, v_L measures the direct effect of unskilled worker employment in the T sector on T sector productivity growth, and v_S measures the direct effect of skilled worker employment in the T sector on T sector productivity growth. The constants δ_T and δ_N measure, respectively, the size of spillover effects of LBD in the T sector on N sector productivity growth, and of LBD in the N sector on T sector productivity growth. In this general specification, both sectors have the capacity to generate productivity growth, as well as to benefit from productivity growth in the other sector.

In this general specification, it is possible to have balanced growth, with the *ratio* of productivity levels in the two sectors constant, but not necessarily equal to one. If we define this ratio by $\phi = A_T/A_N$, then the rate of change of the ratio over time is governed by

$$\frac{\dot{\phi}}{\phi} = \frac{\dot{A}_T}{A_T} - \frac{\dot{A}_N}{A_N} = -(1-\delta_N)(u_L\lambda_{LN} + u_S\lambda_{SN}) + (1-\delta_T)(v_L\lambda_{LT} + v_S\lambda_{ST}). \tag{10}$$

To determine dynamic stability, it is necessary to see for what range of ϕ the growth rate of the productivity differential, $\dot{\phi}/\phi$, is positive and for what range it is negative. Let ϕ^* denote the value at which $\dot{\phi}/\phi = 0$. To achieve such a dynamic equilibrium, we solve the equation implied by $\dot{\phi}/\phi = 0$ to obtain the following linear relationship between λ^*_{LN} and λ^*_{SN}, the equilibrium proportion of unskilled and skilled labor in the N sector, respectively:

$$\lambda^*_{SN} = \frac{(\frac{1-\delta_T}{1-\delta_N})(v_L + v_S)}{u_S + (\frac{1-\delta_T}{1-\delta_N})v_S} - \left(\frac{u_L + (\frac{1-\delta_T}{1-\delta_N})v_L}{u_S + (\frac{1-\delta_T}{1-\delta_N})v_S}\right)\lambda^*_{LN} \tag{11}$$

Here we have made use of the fact that both the proportions of skilled labor and the proportions of unskilled labor in the two sectors must sum to one. It should be remarked that, while this relationship between the shares of unskilled and skilled labor in the N sector must always hold in a long run balanced growth equilibrium, shocks to resource income R will induce short run deviations of the labor allocations away from their long run equilibrium values.

It can be seen that the level of ϕ^* corresponding to λ_{LN}^* and λ_{SN}^* depends on the resource income R. If $\dot{\phi}/\phi > 0$ for $\phi > \phi^*$, the system will be unstable, and will exhibit unbalanced growth leading to complete specialization in one sector or the other. Conversely, if $\dot{\phi}/\phi < 0$ for $\phi > \phi^*$, the system will be stable, and will exhibit balanced growth in equilibrium, with the ratio of sector productivities given by ϕ^*. It is the latter case, of course, that most interests us here. The following Proposition states the condition for dynamic stability, along with the factor allocations and the factor income inequality I that obtain in the unique long run dynamically stable equilibrium.

Proposition 5 *Dynamic stability of the system requires that $\sigma \leq 1$. For the case $\sigma < 1$, there exists a unique set of factor allocations in the N and T sectors. These are given by solving*

$$\frac{\theta_{ST}}{\theta_{LT}}\left(\frac{1-\lambda_{LN}^*}{1-\lambda_{SN}^*(\lambda_{LN}^*)}\right) = \frac{\theta_{SN}}{\theta_{LN}}\frac{\lambda_{LN}^*}{\lambda_{SN}^*(\lambda_{LN}^*)} \tag{12}$$

for λ_{LN}^*, where we can write $\lambda_{SN}^* = \lambda_{SN}^*(\lambda_{LN}^*)$ according to equation 11, and λ_{LT}^* and λ_{ST}^* are obtained by invoking full factor utilization. Equilibrium factor inequality is given by $I = \frac{\theta_{SN}}{\theta_{LN}}\frac{\lambda_{LN}^*}{\lambda_{SN}^*}$.

Proof. The proof of this Proposition is contained in the Appendix. ■

Figure 1 illustrates the long run equilibrium in the Edgeworth-Bowley box, with skilled labor on the vertical and unskilled labor on the horizontal axis. The nontraded sector unskilled labor force and nontraded sector skilled labor force are measured as distances from the point O_N in the bottom left hand corner, and the traded sector unskilled and skilled labor forces are measured as distances from the top right hand corner. The curve is drawn in $(\lambda_{LN}, \lambda_{SN})$ space. We have drawn the contract curve so that it lies everywhere below the diagonal 45 degree line $O_N O_T$, which corresponds to the assumption that the N sector is relatively unskilled labor intensive. Also intersected with the contract curve is a bold, downward sloping line, which corresponds to the condition of long run dynamic equilibrium in the general case summarized by equation 11, with parameters chosen so that the long run equilibrium locus is described by the line $\lambda_{SN}^* = 1.5 - 2\lambda_{LN}^*$. More unequal economies will be those in which more activity is concentrated in the relatively skilled labor intensive sector. If this is the T sector, that means that inequality is decreasing as we move right on the horizontal axis from the point O_N: higher values for λ^*_{LN} correspond to lower equilibrium inequality in a balanced growth economy. [4]

Now let us examine the effect of a resource boom, meaning a one-off increase in R, on the path taken by inequality. The immediate shift in factors of production to ensure market clearing will initially imply an increase in the growth rate of A_N and a decrease in the growth rate of A_T relative to the (common) rates of growth in the sectors that obtained before the boom, so that $\lambda_{LN} > \lambda^*_{LN}$ and $\dot{\phi}/\phi < 0$. In the transition to the long run the relative productivity ϕ will continue to fall, and λ_{LN} decrease, until $\lambda_{LN} = \lambda^*_{LN}$ is again reached and a lower equilibrium level of ϕ is obtained. From Proposition 5 and Lemma 1, the long run level of I and therefore G is independent of R. Resource income inequality G^R is independent of R by assumption. We can thus invoke Lemma 2 to conclude that the long run effect of resource booms on total income inequality G^T works solely through the induced increase in the proportion of resource income in total income, in case 1, as well as through changes in the correctional term in cases 2 and 3. The following proposition summarizes these results.

Proposition 6 *Under conditions of balanced growth, a permanent increase in resource income R will have no long run effect on G. The long run effect on G^T is signed as follows: in cases 1 and 3 of Lemma 2, G^T will fall (rise) if $G^R < G$ ($G^R > G$). In case 2, $G^R < G$ is a sufficient, but not a necessary, condition to produce a fall in G^T.*

Given the variation in the relative degrees of non-resource and resource income inequality across countries, as well as the fact that a large proportion of resource income relative to total income is required to induce substantial changes in the long run value of G^T following resource booms, we expect to find little or no change in long run income inequality in response to resource booms.

[4] A natural question our model allows us to examine is how the structure of LBD effects determines the level of long run equilibrium inequality in the balanced growth economy. An analysis of the issue, omitted due to space limitations, is available from the authors upon request.

3 The Empirical Analysis

We now turn to the empirical analysis in which we test the theoretical prediction that income inequality will fall in the short run immediately after a resource boom, and will then increase steadily over time until the initial impact of the boom disappears. In this section, we describe the methodology and the data. Table 1a reports summary statistics. The empirical literature on natural resources and inequality predominantly relies on cross-sectional regressions, which suffer from acknowledged limitations. They are unable to identify dynamics and are therefore particularly ill-suited for testing the short- and long-run effects of resource booms on inequality. They also suffer from omitted variable bias and it is therefore "crucial to move from cross-country to panel data evidence" (Van der Ploeg, 2006). For these reasons, we employ a panel data estimator with country fixed effects and regional time dummies to reduce concerns of omitted variables. We exploit variation in the world prices of a country's commodity exports to identify resource booms, as world prices are typically unaffected by individual countries and are therefore likely to be exogenous (Deaton and Miller, 1995).[5] In particular, we analyze the effects of commodity export prices on income inequality using the following dynamic panel data model in error-correction form[6]:

$$\Delta G_{i,t} = \alpha_i + \delta' z_{i,t} + \lambda G_{i,t-1} + \beta_1' x_{i,t-1} + \beta_2 \Delta G_{i,t-1} + \beta_3' \Delta x_{i,t} + \epsilon_{i,t} \tag{13}$$

In the above equation, the subscripts $i = 1, ...N$ and $t = 1, ...T$ index the countries and years in the panel, respectively. Here $G_{i,t}$ stands for household income inequality in country i in year t, α_i is a country-specific fixed effect, and $z_{i,t}$ is an $rT \times 1$ vector of regional time dummies, where r is the number of regions.[7] The term $x_{i,t-1}$ is an $m \times 1$ vector of m variables that are expected to affect inequality in the short and/or long run. Our dataset includes all countries and years for

[5] Since some major commodity-exporting countries may have an influence on world prices, we investigate the robustness of our results to the exclusion of major exporters as part of our sensitivity analysis.

[6] The error-correction form is a convenient reparameterization of an autoregressive distributed lag model with 2 lagged levels of the dependent variable and the contemporaneous and first lagged level of the independent variables.

[7] The regional time dummies capture year fixed effects for the following regions: Central and Eastern Europe and Central Asia, East Asia and Pacific and Oceania, Latin America and Caribbean, North Africa and Middle East, South Asia, Sub-Saharan Africa, and Western Europe and North-America. This categorization is based on the classifications of the World Bank and the United Nations, and on the Central and Eastern European Directory.

which data are available, and covers 90 countries (listed in Table 1b) for the period 1965 to 1999. We next discuss how the key components of equation (13) were constructed.

3.1 Income inequality and commodity export prices

Our measure of household income inequality, $G_{i,t}$, is the Gini index constructed by Galbraith and Kum (2005) and is based on the Deininger and Squire (1996) household income inequality data set and the UTIP-UNIDO data set on manufacturing pay inequality.[8] It is much more comprehensive than the Deininger and Squire inequality measure and uses a more informed filling-in of missing observations than other studies. It should be noted that the use of existing cross-country data sets on income inequality is not without problems. Atkinson and Brandolini (2001), for example, point at important cross-country differences in definitions of inequality and differences in the data used to construct measures of inequality, which can affect the estimated inequality levels as well as trends in inequality over time. Galbraith and Kum (2005) argue that their data set addresses some of these problems and provides a more comparable and consistent measure of cross-country household income inequality than the Deininger and Squire (1996) data set.

The vector $x_{i,t-1}$ includes a commodity export price index, which was constructed using the methodology of Deaton and Miller (1995). We first collect data on world commodity prices and commodity exports for as many commodities as data availability allowed.[9] We then construct weights by dividing the individual 1990 export values for each commodity by the total value of 1990 commodity exports for each country. These weights are held fixed over time and applied to the world price indices of the same commodities to form a country-specific geometrically weighted index of commodity export prices.[10] This index was first constructed on a quarterly basis and

[8] We use the variable "EHII2.3", see Galbraith and Kum (2005) for details.

[9] Prices are from the IMF's International Financial Statistics (IFS) and the Energy Information Administration. Exports are from the UNCTAD Commodity Yearbook and the UN International Trade Statistics Yearbook.

[10] Following Deaton and Miller (1995), the weights are held fixed in order to construct an exogenous index that does not include endogenous supply responses to world prices. This means that we lose some important changes in the composition of primary exports but, as recognized by Deaton and Miller (1995), this loss is inevitable if we are to exclude endogenous quantity changes. Moreoever, the loss is likely to be limited as the pairwise correlations

deflated by the export unit value (IFS).[11] We then constructed the annual index by taking the log of the annual average (rescaled so that 1980 = 100) of the quarterly index. Finally, to allow the effect of commodity prices to be larger for more commodity-dependent countries, we weight the annual index by the share of commodity exports in a country's GDP.[12] In addition to this general index, we also constructed separate indices for non-agricultural and agricultural commodities.[13]

Table 1b lists the shares of commodity exports in GDP. While the median share is 0.05, there is considerable cross-country variation in the relative importance of natural resources. Some countries hardly export resources, while others' export revenues correspond to major shares of GDP (for example Zambia and Venezuela). Despite the substantial coverage of the Galbraith and Kum (2005) inequality data set, some large resource exporters like Nigeria, Oman and Saudi Arabia are not included in our sample, mostly due to lack of data on income inequality.

3.2 Long run determinants of income inequality

The theoretical literature on income inequality suggests several determinants. First, inequality has often been related to per capita income. Kuznets (1955) considered an economy with an agricultural sector and an industrial sector. Initially, the majority of workers are employed in agriculture. As an economy develops, workers move from low-income jobs in agriculture to higher-income jobs in the industrial sector, which given the relative size of both sectors initially increases income inequality. As more workers move to the high-income industrial sector and more workers

between the 1990 weights and the same weights for 1970, 1980, and 2000, are 0.74, 0.87, and 0.84, respectively, indicating that the weights of individual resources in a country's primary exports are relatively persistent.

[11] Short gaps in a few price series were filled by inter- or extrapolation. Where gaps for unimportant commodities (share in total exports < 10% or share in GDP < 1%) would cause missing values, these price series were excluded.

[12] We relax the assumption that the price effect *linearly* increases with the share of exports when we discuss the estimation results. GDP in current US dollars is from the World Bank's World Development Indicators (WDI).

[13] Our sample includes 15 non-agricultural commodities (aluminum, coal, copper, gasoline, ironore, lead, natural gas, nickel, oil, phosphatrock, silver, tin, uranium, urea, zinc) and 35 agricultural commodities (bananas, barley, butter, cocoabeans, coconutoil, coffee, copra, cotton, fish, groundnutoil, groundnuts, hides, jute, maize, oliveoil, oranges, palmkerneloil, palmoil, pepper, plywood, poultry, pulp, rice, rubber, sisal, sorghum, soybeanoil, soybeans, sugar, sunfloweroil, swinemeat, tea, tobacco, wheat, wool).

increase their income *within* that sector, inequality starts to fall. In addition, fewer workers in agriculture means a higher relative wage in that sector, which further adds to the fall in inequality. Empirical support for the Kuznets curve was found by Ahluwalia (1976) and Barro (2000).

Inequality is also often associated with education. Tinbergen (1975) showed how higher levels of education increase the supply of skilled labour, which lowers relative wages and wage inequality. In Saint-Paul and Verdier (1993), public education is used for intra-generational income redistribution. Initially, inequality is high and the poor median voter prefers a high level of redistribution through public education. Over time, the increase in human capital through education lowers inequality.

Finally, Acemoglu and Robinson (2002) argue that *political* rather than *economic* factors are crucial to understanding inequality. They explain how industrialization in the West increased inequality but also mobilized the poor by concentrating them in urban centers and factories. This led to political instability or a threat of revolution, forcing democratization on political elites, which led to institutional changes that encouraged redistribution and lowered inequality.

Following the theoretical literature, as well as the recent empirical work by Barro (2000), we include three long run control variables in the vector $x_{i,t-1}$: log real GDP per capita, a measure of democracy (based on the number of political constraints), and a measure of educational attainment (expressed as the average years of primary schooling of the population aged 15 and over).[14] These control variables are statistically significant in the specification of equation (13).[15]

[14] GDP per capita in constant 2000 US$ is from WDI, the political constraints indicator is the variable "POLCONV" from Henisz (2000), and the education variable is from Barro and Lee (2000) and was linearly interpolated.

[15] We considered a wide range of additional controls, including GDP per capita squared (to test the hypothesis underlying the Kuznets curve), measures of secondary and higher schooling, alternative measures of democracy, various governance indicators, measures of trade and capital account openness, inflation, external debt, political violence, commodity price volatility, measures of financial and industrial development, natural disasters, and the political orientation of government parties or the executive. These variables are not included in our preferred specification because they were either not robustly significant or severely lowered the number of observations. However, we do use them when we test the robustness of our results in the next section.

3.3 Testing for the existence of a long run relationship

The error-correction model in equation (13) is only appropriate if there is a long-run relationship between income inequality and GDP per capita, democracy, and schooling.[16] Testing for the existence of such a relationship is usually done using cointegration techniques. Cointegration requires that the individual variables are integrated of order 1, and that the residuals of a regression of $G_{i,t}$ on $x_{i,t}$ are stationary. To test this, we performed panel unit root tests on the levels and the differences of the variables in $G_{i,t}$ and $x_{i,t}$ and then performed a panel cointegration test.[17] The results of these tests, available upon request from the authors, indicate that the variables are cointegrated and that the error-correction specification in equation (13) is valid.

A potential weakness of the cointegration techniques is that they require the variables to be non-stationary. Although this assumption can be tested using unit root tests, as we have done above, such tests are not without problems and introduce a further degree of uncertainty into the analysis of levels relationships (Pesaran, Shin, and Smith, 2001). In addition, on theoretical grounds, it is not obvious that income inequality should be non-stationary and cointegrated with variables such as GDP per capita and primary schooling. Pesaran, Shin, and Smith (2001) propose a new approach to testing the existence of a long-run relationship which is applicable irrespective of whether the underlying regressors are stationary or non-stationary. Their "bounds tests" are based on a standard F-statistic for the null hypothesis that the coefficients of the lagged level variables in the error-correction model are equal to zero. The asymptotic distribution of the F-statistic is non-standard under the null of no long-run relationship. Pesaran, Shin and Smith (2001) provide two sets of critical values: One set assuming that the lagged level variables are all non-stationary, and another set assuming that they are all stationary. Evaluating the F-statistic against the two relevant critical values from both sets, results in three possible outcomes. If the F-statistic lies

[16] Commodity export prices could also be part of this long-run relationship but, as will become evident below, we do not find evidence of a long-run effect of commodity export prices on inequality.

[17] We use the Im, Pesaran and Shin (2003) panel unit root test for a balanced sub-sample of 17 countries and 35 years, the Maddala and Wu (1999) panel unit root test for both the full sample and the balanced sub-sample, and a panel cointegration test by Pedroni (1999).

below the two relevant critical values, the null hypothesis cannot be rejected, regardless of whether the variables are stationary or non-stationary. If the F-statistic lies *in between* the two relevant critical values, the result is inconclusive and rejection of the null depends on whether the variables are stationary or non-stationary. Finally, if the F-statistic *exceeds* the two relevant critical values, the null hypothesis is always rejected, regardless of whether the variables are stationary or non-stationary. This last possibility is especially of interest in this context, as it confirms the existence of a long-run relationship irrespective of the order of integration of the variables. Hence, even if one doubts the conclusiveness of the panel unit root and panel cointegration tests, one can still draw inference from the error-correction model in equation (13), as the long-run relationship assumed in the model exists irrespective of whether the variables are I(1) or I(0). Following Pesaran, Shin and Smith (2001), we calculated the F-statistic and evaluated it against the two relevant non-standard critical values corresponding to the 1% significance level. The value of the F-statistic (10.84) exceeded the the two critical values (5.17 and 6.36) and hence we reject the null of no long-run relationship at the 1% level. This is reassuring as it confirms the existence of a long-run relationship between income inequality and GDP per capita, political constraints and primary schooling, regardless of whether the variables are stationary or non-stationary.

4 Estimation Results

The results of estimating equation (13) are reported in Table 2.[18] Columns (1) and (2) show the OLS and fixed effects results for the baseline specification without the commodity export price variables. The long-run variables enter with the expected signs but are only significant in the specification with country fixed effects, suggesting that the inclusion of fixed effects is indeed important.[19] The lagged level of inequality enters with a negative sign and is significant at 1%. The size of the coefficient in the (preferred) fixed effects specification indicates that the speed of

[18] The long-run coefficients correspond to $-(\frac{1}{\lambda}) \cdot \beta_1$, while the short-run coefficients correspond to λ, β_2, and β_3.

[19] We tested the Kuznets relation by adding GDP per capita squared but the coefficient was insignificant. This may reflect that the relation has weakened over time (Anand and Kanbur, 1993) or that it works better for a cross section of countries in a given year than for inequality within countries over time (Li, Squire and Zou, 1998).

adjustment to long-run equilibrium is around 20% per year. The lagged change in inequality also enters with a negative sign and is significant. We experimented with additional lags, as well as with lags of the changes in the long-run variables, but found these to be unimportant.

In Table 2, column (3), we include the lagged level and the change in the commodity export price index to test the long-run and short-run effects of commodity export prices on inequality. The long-run coefficient is positive but highly insignificant, indicating that commodity export prices do not affect the long-run level of inequality.[20] The short-run coefficient is negative, suggesting that commodity booms lower inequality in the same year. However, the coefficient is not significant so should be viewed with caution.[21] In Table 2, column (4), we drop the lagged level of the commodity export price index and find a similar result for the change in the index.

We next investigate whether the short-run effect of higher commodity prices varies across non-agricultural and agricultural commodities by replacing the change in the general index with the changes in the sub-indices for both types of commodities. As we explain below, this distinction may be important as the revenues from non-agricultural exports typically accrue to governments, whereas the revenues from agricultural exports accrue predominantly to farmer households. The

[20] Collier and Goderis (2008) show that commodity booms have adverse long-term effects on GDP per capita. We investigated the possibility that commodity export prices affect long-run inequality *indirectly* through GDP per capita by rerunning the specification in Table 2, column (3), for the same sample but without GDP per capita. The commodity export price index now entered with a negative sign but the coefficient remained highly insignificant with a p-value of 0.95, suggesting that any indirect effect of commodity prices through GDP per capita is likely to be negligible. As an alternative way of testing the long-run relation between natural resources and income inequality, we also ran cross-sectional OLS regressions of the level of inequality on commodity exports over GDP and several controls, all in 1990 (see Alexeev and Conrad, 2009, for a similar approach to testing the relation between natural resources and long-run *economic growth*). Using 10 alternative sets of controls from both our panel analysis and Barro (2000), we do not find a significant effect of natural resources on the long-run level of inequality, consistent with our panel data results. For robustness, we also ran the regressions with inequality in 1999 (the last available year) to allow any effect to take more time, and with inequality in 1999 *and* independent variables in 2000. In both cases, we again do not find a robustly significant effect of natural resources on long-run inequality.

[21] We experimented with lags of the change in the commodity export price index but found these to be unimportant. We also allowed for an effect of higher oil prices on oil *importing* countries by including an oil import price index, but found no evidence of a systematic effect. All our results go through when including this oil import price index.

results are reported in Table 2, column (5). Both variables enter with a negative sign but the change in the non-agricultural export price index is now significant at 5%. This indicates that a rise in the prices of non-agricultural commodities lowers inequality in the same year. The change in the agricultural index is not significant but its coefficient is only slightly smaller than the coefficient for the change in the non-agricultural index. In fact, an F-test did not reject the null hypothesis of equal coefficients with a p-value of 0.96.[22] A possible reason for why we only find a statistically significant effect for non-agricultural exports is that the sample variation of the change in the agricultural index is smaller than the variation of the change in the non-agricultural index. The standard error of the coefficient is therefore larger and the significance lower. This could explain why the estimated coefficients for non-agricultural and agricultural exports are similar in size and not significantly different, while the standard error of the latter coefficient is much higher. But the difference in significance may also reflect a genuine difference in spending patterns. Non-agricultural revenues typically end up with governments, which may spend a relatively large part of it, whereas revenues from agriculture accrue mostly to farmers, who may save more. Robinson, Torvik and Verdier (2006) argue that politicians discount the future by the probability of being in power and therefore over-extract natural resources. The same argument could be used to explain why politicians consume a larger part of the revenues out of resource booms. A recent study by the World Bank (2006) indeed indicates that countries with a large share of mineral and energy (non-agricultural) rents in GNI typically have lower genuine saving rates. Using data from this study, we ran a cross-sectional OLS regression of the genuine saving rate on the share of non-agricultural commodity exports in GDP from our dataset. Consistent with the findings of the World Bank, we find a negative coefficient, significant at 1%, suggesting that countries with a

[22] Since its long-run coefficient was insignificant in Table 2, column (3), we exclude the lagged *level* of the commodity export price index from the specifications of Table 2, columns (5) and (6), and Table 3, columns (1) to (4). However, all results on the short-run effects of non-agricultural and agricultural export prices are robust to the inclusion of the lagged level of the commodity export price index or the lagged level of the non-agricultural index. When included, their long-run coefficients are always insignificant, consistent with our earlier finding that commodity prices do not affect long-run inequality. We also experimented with the agricultural index but again found no significant long-run effect.

large share of non-agricultural exports on average have lower genuine saving rates. We then ran the same regression for agricultural exports and now found a positive but insignificant coefficient, suggesting that, in contrast to non-agricultural exporters, *agricultural* resource exporters do *not* have lower genuine saving rates.[23] Although these correlations should not be interpreted as causal, they are consistent with the hypothesis that non-agricultural resource booms have stronger Dutch disease effects because governments consume a larger part of the revenues from booms than farmers. In Table 2, column (6), we exclude the insignificant effect of agricultural prices and find almost identical results. These results are consistent with our theoretical prediction that inequality should fall in the short run in response to a resource boom.

We next use the results in Table 2, column (6), to graphically show the time path of inequality following a resource boom. Figure 2 shows three impulse response functions of inequality for a two standard deviations (33 % points) increase in the growth rate of non-agricultural export prices for countries with different levels of non-agricultural exports. The estimated declines in inequality in the year of the boom are 0.11, 0.22, and 0.32 Gini index points for countries with 10%, 20%, or 30% non-agricultural exports to GDP, respectively.[24] After its decline in the year of the boom, inequality gradually moves back to its original pre-boom level. The speed of adjustment is such that around five years after the shock, two thirds of its initial impact has died out. These results on the time path of inequality after a boom lend support to the predictions from the theory.

It should be noted that, while the effect of resource booms is important for *resource-rich* countries, it is (unsurprisingly) much less so for other countries. Recall that we allow for this in our estimations by weighting the commodity export price indices by a country's share of commodity exports in GDP.[25] Hence, for the median country in our sample with a share of non-agricultural

[23] Including both non-agricultural and agricultural exports in one regression yielded almost identical results.

[24] These estimates were calculated by multiplying the coefficient in Table 2, column (6), by the share of non-agricultural exports in GDP and the two standard deviations price change (e.g. $-3.26 * 0.10 * 0.33 = -0.11$).

[25] This assumes that the importance of price effects linearly increases with the share of exports. Since price effects may only matter for countries with high export levels, we re-estimated the specifications in Table 2, but decomposing the effects of commodity export prices into the effects for the top quintile exporters and the effects for the other sample countries. The coefficients for the top quintile exporters were not significantly different from the coefficients

exports in GDP of only 0.01, the estimated short-run decline in inequality from a two standard deviations increase in the growth rate of non-agricultural prices is only $-3.26 * 0.01 * 0.33 = 0.01$ Gini index points. But for resource-rich countries with shares of 0.09 (80^{th} percentile of sample distribution), 0.14 (90^{th} percentile), 0.18 (95^{th} percentile), and 0.34 (99^{th} percentile), the estimated declines in inequality are 0.10, 0.15, 0.19 and 0.37 Gini index points, respectively.

We next test the robustness of our finding that higher non-agricultural commodity prices lower short-run inequality. Table 3, column (1), shows the results when excluding the regional time dummies. The short-run effect of non-agricultural export prices remains significant at 5%, while the size of the coefficient is somewhat smaller. In Table 3, column (2), we decompose the change in the non-agricultural index into positive and negative changes. The coefficient for the *positive* shocks is negative, significant at 5%, and larger than the coefficient in earlier specifications, confirming that resource booms lower inequality in the year of the boom. Although the coefficient for the *negative* shocks is insignificant, it is also negative and is not significantly different from the coefficient for positive shocks, hence suggesting that pooling positive and negative shocks is appropriate. This is consistent with our theory, which predicts that the short run response of the Gini to a resource bust would be opposite in sign, but generally similar in magnitude, to the response to a boom. Table 3, column (3), shows that our results are also robust to using OLS instead of fixed effects.

Although world commodity prices are typically unaffected by individual countries and our indices exclude endogenous supply responses, our estimates could suffer from endogeneity if major commodity exporters have an influence on world prices. To address this concern, we express individual countries' exports of each commodity as a share of world exports. We find that of the 90 countries in our sample, 22 countries export at least one commodity for which their share in world exports exceeds 20%, while 34 countries export at least one commodity for which their share in world exports exceeds 10%. We investigate whether excluding these major exporters affects our

for the other countries. We also tested whether price effects only occur in case of large shocks by adding squared terms of the price indices to the specifications in Table 2, columns (3) to (6), but the coefficients of the squared terms were not significant. The absence of non-linearities beyond the linearly increasing importance of price effects suggests that the weighting of the indices by a country's share of exports in GDP is appropriate.

results by re-estimating the specification in Table 2, column (6), for a sub-sample without the 22 countries with exports larger than 20% and a sub-sample without the 34 countries with exports larger than 10%. Although this decreases the sample size by almost a third and half, respectively, the coefficient of the change in the non-agricultural price index only marginally changes from -3.26 (5% significance) to -3.30 (10% significance) and -3.76 (10% significance) for the two sub-samples, respectively. Hence, our results do not seem to be biased by major commodity exporters.

Endogeneity could also arise from weighting the commodity export price indices by a country's share of commodity exports in GDP, as this share may be driven by time-varying factors that also affect inequality, such as government policies. Resource-rich countries with bad policies that cause low growth and high inequality, for example, also tend to have high shares of commodity exports in GDP due to the lack of development of their non-resource sectors. This implies that we attach high weights to poor countries, which could bias the results. To address this concern, we instrument for the share of non-agricultural exports in GDP, using estimates of the per capita stock of sub-soil assets in 2000 from the World Bank (2006) as an instrument. These estimates are based on the net present value of a country's expected benefits over a horizon of 20 years and include 13 commodities, 12 of which are also included in our non-agricultural index. Since the share of non-agricultural exports in GDP only enters our specifications as a weight of the non-agricultural export price index, we create an instrument by reconstructing the index but weighting it by the value of per capita sub-soil assets instead of the share of non-agricultural exports in GDP. We then use the change in this variable as an instrument for the change in the non-agricultural export price index in the specification of Table 2, column (6). For this instrument to be valid, it should be correlated with the (potentially) endogenous regressor and not be correlated with the error term. The first requirement is likely to be fulfilled, as a country's level of commodity exports is correlated with its natural wealth. The second requirement is less likely to be fulfilled, as countries with high inequality tend to be poorer and have smaller resource stocks due to less geological exploration and more overexploitation. Weighting by the value of sub-soil assets may therefore imply giving high weights to rich countries. Although this could bias the results, the direction of the bias is likely to be opposite to the bias in the uninstrumented regressions, where high weights were given to poor countries. Comparing the coefficients of the instrumented and uninstrumented regressions can therefore shed light on the size of the potential bias and the numerical range within which the actual coefficient is likely to be located. The second-stage IV results are reported in Table 3, column (4).[26] The change in the non-agricultural export index again enters with a negative sign and, although the coefficient is no longer significant, it is somewhat larger than in Table 2, column (6). Given that any biases in the OLS and IV estimates are likely to have opposite signs, this suggests that the bias in the OLS estimates is likely to be small and, if anything, leads to an underestimation of the short-run effect of non-agricultural export prices. In fact, a Davidson-MacKinnon test of exogeneity did not reject the null of exogeneity with a p-value of 0.71.[27]

Finally, we performed 50 other robustness checks by adding additional controls to the specification in Table 2, column (6). The variables we consider are GDP squared, alternative measures of schooling, alternative indicators of democracy, several indicators of institutional quality, several indicators that capture the type of government (left wing, centre, or right wing), and a range of variables from the empirical growth literature. The results confirmed our finding that higher non-agricultural export prices lower inequality in the same year.[28]

[26] The instrument enters with the expected positive sign and is significant at 1% in the first stage.

[27] For robustness, we repeated the IV estimation using the 1994 estimates of sub-soil assets instead of the 2000 ones and found similar results. We also performed the IV estimation for the two sub-samples in which we exclude the major commodity exporters. The coefficients for the change in the non-agricultural index were similar or larger: -3.94 (2000 sub-soil assets) and -4.05 (1994 sub-soil assets) for the sub-sample without the countries with exports > 20 %, and $-7.17^{**}$and -6.81^{**} for the sub-sample without the countries with exports > 10 %.

[28] For each additional control, we ran a regression with the lagged level and a regression with the lagged level, contemporaneous difference, and any additional significant lagged differences. The coefficient of the change in the non-agricultural index had a mean value of -3.43 and varied between -5.29 and -2.15, which is close to the value in Table 2, column (6). It was significant at 5% in 32 cases, significant at 10% in 11 other cases, and insignificant in 7 cases. In 6 of these last 7 cases, additional estimations revealed that the coefficient turned insignificant due to the smaller sample and not because of the additional control. We also ran two regressions including all additional controls together. Despite the substantially smaller sample, the coefficient remained significant at 10%.

4.1 Commodity price volatility and uncertainty

We have so far ignored an important alternative channel through which natural resources may affect long-run inequality: the *volatility* of commodity prices. Volatility reduces the collateral value of inventories which increases borrowing costs, especially for poor credit constrained firms or farm households. In addition, volatility increases the incidence of defaults among the poor, as they typically lack savings and access to liquidity to deal with adverse price shocks. But it could also raise inequality by forcing poor farmers to diversify their crops, which reduces their benefits from specialization (Dehn, 2000). And finally, volatility can increase inequality through lowering investments by poor risk-averse investors (Servén, 1998, Zeira, 1990).

If volatility indeed affects long-run inequality and is correlated with our commodity export price index, our estimate of the long-run effect of commodity prices may (partly) reflect the effect of the *volatility* of prices rather than the *level*. To investigate this possibility, we constructed a measure of volatility by taking the pre-1983 (median sample year) mean absolute change in the unweighted commodity export price index (multiplied by 100) for the years until 1983 and the post-1982 mean absolute change for the years from 1983. We added the first lag of this variable, weighted by commodity exports over GDP, to the specification of Table 2, column (3), in which we estimated the long-run effect of commodity export prices. The long-run coefficient of commodity export prices changed from 0.48 to -0.14 but remained insignificant (p-value = 0.98). The long-run coefficient of volatility had the expected positive sign, indicating that volatility leads to higher long-run inequality, but the coefficient was insignificant (p-value = 0.23).[29]

The volatility measure described above reflects *ex post volatility* in the sense that it captures *realized* changes in commodity prices. However, *ex ante uncertainty* about prices may be more important for inequality than realized volatility. Moreover, ex post volatility includes all price

[29] We found the same coefficient for volatility when excluding the lagged level of the commodity price index. For robustness, we constructed a second measure of volatility by calculating for each quarter the country-specific standard deviation of the quarterly commodity export price index from section 3.1 over the quarter and the three quarters preceding it, and then taking the log of the annual average of this "rolling standard deviation", weighted by exports over GDP. Using this measure, we found the same results as for the first volatility measure.

movements, while ex ante uncertainty should only include the unpredictable component of price changes (Ramey and Ramey, 1995, Servén, 1998). We therefore also experimented with a measure of export price uncertainty. Following Dehn (2000) and Servén (1998), we estimate for each country separately the following generalized autoregressive conditional heteroskedasticity (GARCH (1,1)) model in which actual price volatility is explained by past volatility and past expected volatility:

$$\begin{aligned} \Delta I_t &= \alpha_0 + \alpha_1 t + \beta_1 \Delta I_{t-1} + \beta_2 I_{t-2} + \beta_3 D_t + \varepsilon_t \\ \sigma_t^2 &= \gamma_0 + \gamma_1 \varepsilon_{t-1}^2 + \gamma_2 \sigma_{t-1}^2 \end{aligned} \tag{14}$$

where I_t is the unweighted log commodity export price index in quarter t, t is a linear time trend, D_t is a vector of quarterly dummies to remove seasonal effects, and σ_t^2 denotes the variance of ε_t, conditional upon information up to period t. We use the annual average of the fitted values of the second equation in (14), weighted by the % share of exports in GDP, as a measure of commodity export price uncertainty, since it captures the predicted variance of the changes in commodity prices from past actual and expected volatility. We add this measure of uncertainty to the specification of Table 2, column (3). The results are reported in Table 3, column (5). The long-run coefficient for commodity export prices remains insignificant, but the long-run effect of commodity export price uncertainty on income inequality is positive and significant at 1%. This indicates that it is the ex ante uncertainty about export prices that matters significantly for long-run income inequality. For robustness, we reran the same specification without the lagged level of the commodity export price index and found an almost identical long-run coefficient for volatility (results reported in Table 3, column (6)). For a country like the Republic of Congo with commodity exports of 29% of GDP, the results in Table 3, column (5), imply that an increase of two standard deviations in the predicted volatility of export prices (the annual average of the fitted values of the second equation in (14)) leads to a $1.33 \times 29 \times 0.028 = 1.1$ point increase in the long-run Gini index of inequality.

Summarizing, our estimates of the long-run effect of commodity prices on income inequality do not seem to be explained by price volatility or uncertainty. When controlling for volatility or uncertainty, we still find an insignificant long-run effect. However, our results do indicate that uncertainty about commodity export prices significantly increases long-run inequality, consistent with the notion that the poor are less able to deal with booms and busts than the rich.

5 Conclusions

This paper has theoretically and empirically analyzed the time path of income inequality following a natural resource boom. Our main finding is that resource booms, especially oil and mineral booms, lower inequality in the year of the boom. This effect then gradually diminishes over time until inequality returns to its pre-boom level in the long run. Our complementary finding that commodity price uncertainty increases long-run income inequality is of independent interest and provides an interesting avenue for future research.

Two comments are in order. First, although our sample includes 90 countries, some large resource exporters like Nigeria, Oman, and Saudi Arabia, are not included, mostly due to lack of data on income inequality. As more inequality data become available, it may be worthwhile to investigate whether our findings extend to these countries. Second, this paper has ignored some alternative mechanisms through which booms may affect inequality. One example is public sector employment. Resource windfalls often generate new government jobs, which may reduce income inequality (Ross, 2007). However, such jobs are often created to buy political support and may come at a cost of lower growth (Robinson, Torvik and Verdier, 2006).

Appendix: Proof of Proposition 5

To derive the condition for dynamic stability, we must determine the sign of the rate of change of relative productivity growth. This is found by computing

$$\frac{d(\dot{\phi}/\phi)}{dt} = -\left([u_L(1-\delta_N) + v_L(1-\delta_T)] \frac{d\lambda_{LN}}{d\phi} + [u_S(1-\delta_N) + v_S(1-\delta_T)] \frac{d\lambda_{SN}}{d\phi} \right) \frac{d\phi}{dt} \qquad (15)$$

For stability, we need $\frac{d(\dot{\phi}/\phi)}{dt}$ to take the opposite sign from $\frac{d\phi}{dt}$. This condition is satisfied if and only if the static effect of increased ϕ on the linear combination of both types of employment in the N sector, which is given by the large term in parenthesis in the above expression, is positive. Since $\frac{d\lambda_{LN}}{d\phi}$ and $\frac{d\lambda_{SN}}{d\phi}$ take the same sign always, we can reduce the problem of determining dynamic stability to the problem of determining necessary and sufficient conditions for $\frac{d\lambda_{LN}}{d\phi} > 0$. This amounts to the same thing as in Torvik (2001): dynamic stability requires that $\sigma \leq 1$. The comparative static computations are available from the authors upon request, and the intuition is similar to that in Torvik (2001). Low demand elasticity induces demand-side shifts in consumption patterns that are sufficient to counteract the factor movement effects, due to technical change biased in favor of one sector, that would otherwise result in complete specialization for $\sigma > 1$.

Now let us solve for long run inequality and factor allocations. In equilibrium, the values of λ_{LN} and λ_{SN} are fixed by two conditions. The first is the dynamic stability condition given by equation 11. The second is factor market equilibrium (characterized by factor price equalization between the sectors and profit maximization). Factor market equilibrium defines the *contract curve*, which traces out the locus of pairs $(\lambda_{LN}, \lambda_{SN})$ in the Edgeworth-Bowley box for the economy. For the Cobb-Douglas production functions we have assumed, the contract curve is defined by the relation

$$\frac{\theta_{ST}}{\theta_{LT}} \frac{\lambda_{LT}}{\lambda_{ST}} = \frac{\theta_{SN}}{\theta_{LN}} \frac{\lambda_{LN}}{\lambda_{SN}}. \qquad (16)$$

Both sides of the above equation are equal to our measure of inequality I and are independent of ϕ. Making the substitutions $\lambda_{LT} = 1 - \lambda_{LN}$, and $\lambda_{ST} = 1 - \lambda_{SN}$, which must obtain according to full utilization of unskilled and skilled labor, and using the fact that equation 11 must hold in a long run dynamic equilibrium, it follows that the equilibrium fraction of unskilled labor in the N sector, which we will denote by λ^*_{LN}, is determined by the solution to equation 12.

References

[1] Acemoglu, D. and Robinson, J.A. (2002), The Political Economy of the Kuznets Curve, *Review of Development Economics 6*, 183-203.

[2] Ahluwalia, M. (1976), Income Distribution and Development: Some Stylized Facts, *American Economic Review 66*, 128-135.

[3] Alexeev, M. and Conrad, R. (2009), The Elusive Curse of Oil, *Review of Economics and Statistics 91*.

[4] Anand, S. and Kanbur, R.M. (1993), Inequality and Development: A Critique, *Journal of Development Economics 41*, 19-44.

[5] Atkinson, A. B. and Brandolini, A. (2001), Promise and Pitfalls in the Use of 'Secondary' Datasets: Income Inequality in OECD Countries, *Journal of Economic Literature 39*, 771-799.

[6] Barro, R.J. (2000), Inequality and Growth in a Panel of Countries, *Journal of Economic Growth 5*, 87-120.

[7] Barro, R.J. and Lee, J. (2000), International Data on Educational Attainment: Updates and Implications, CID Working Paper 42.

[8] Collier, P., Goderis, B. (2008), Commodity Prices, Growth and the Natural Resource Curse: Reconciling a Conundrum, OxCarre Research Paper 2008-14.

[9] Corden, W. M. and Neary, J.P. (1982), Booming Sector and De-Industrialization in a Small Open Economy, *The Economic Journal 92*, 825-848.

[10] Daudey, E. and García-Peñalosa, C. (2007), The Personal and Factor Distributions of Income in a Cross-Section of Countries, *Journal of Development Studies 43*, 812-829.

[11] Deaton, A.S. and Miller, R.I. (1995), International Commodity Prices, Macroeconomic Performance, and Politics in Sub-Saharan Africa, *Princeton Studies in International Finance 79*.

[12] Dehn, J. (2000), Commodity Price Uncertainty in Developing Countries, CSAE Working Paper 2000-12.

[13] Deininger, K. and Squire, L. (1996), A New Data Set Measuring Income Inequality, *World Bank Economic Review 10*, 565-591.

[14] Galbraith, J.K. and Kum, H. (2005), Estimating the Inequality of Household Incomes: A Statistical Approach to the Creation of a Dense and Consistent Global Data Set, *Review of Income and Wealth 51*, 115-143.

[15] Gylfason, T. and Zoega, G. (2003), Inequality and Economic Growth: Do Natural Resources Matter?, in Eicher, T. and Turnovsky, S. (Eds.), *Inequality and Growth: Theory and Policy Implications*, MIT Press, Cambridge (MA) and London.

[16] Henisz, W.J. (2000), The Institutional Environment for Economic Growth, *Economics and Politics 12*, 1-31.

[17] Im, K.S., Pesaran, M.H. and Shin, Y. (2003), Testing for Unit Roots in Heterogeneous Panels, *Journal of Econometrics 115*, 53-74.

[18] Kuznets, S. (1955), Economic Growth and Income Inequality, *American Economic Review 65*, 1-28.

[19] Leamer, E.E., Maul, H., Rodriguez, S. and Schott, P.K. (1999), Does Natural Resource Abundance Increase Latin American Income Inequality?, *Journal of Development Economics 59*, 3–42.

[20] Li, H., Squire, L. and Zou, H. (1998), Explaining International and Intertemporal Variations in Income Inequality, *The Economic Journal 108*, 26-43.

[21] Maddala, G.S. and Wu, S. (1999), A Comparative Study of Unit Root Tests With Panel Data and a New Simple Test, *Oxford Bulletin of Economics and Statistics 61*, 631-652.

[22] Obstfeld, M. and Rogoff, K. (1996), *Foundations of International Macroeconomics*, MIT Press, Cambridge (MA).

[23] Pedroni, P. (1999), Critical Values for Cointegration Tests in Heterogeneous Panels with Multiple Regressors, *Oxford Bulletin of Economics and Statistics 61*, 653-670.

[24] Pesaran, M.H., Shin, Y. and Smith, R.J. (2001), Bounds Testing Approaches to the Analysis of Level Relationships, *Journal of Applied Econometrics 16*, 289–326.

[25] Ramey, G. and Ramey, V.A. (1995), Cross-Country Evidence on the Link between Volatility and Growth, *American Economic Review 85*, 1138-1151.

[26] Robinson, J.A., Torvik, R. and Verdier, T. (2006), The Political Foundations of the Resource Curse, *Journal of Development Economics 79*, 447-468.

[27] Ross, M. (2007), How Can Mineral Rich States Reduce Inequality?, in Sachs, J.D., Stiglitz, J.E. and Humphreys, M. (Eds.), *Escaping the Resource Curse*, Columbia University Press, New York.

[28] Saint-Paul, G. and Verdier, T. (1993), Education, Democracy and Growth, *Journal of Development Economics 42*, 399-407.

[29] Servén, L. (1998), Macroeconomic Uncertainty and Private Investment in LDCs: An Empirical Investigation, mimeo World Bank.

[30] Sokoloff, K.L. and Engerman, S.L. (2000), History Lessons: Institutions, Factor Endowments, and Paths of Development in the New World, *Journal of Economic Perspectives 14*, 217-232.

[31] Tinbergen, J. (1975), *Income Distribution*, North Holland, Amsterdam.

[32] Torvik, R. (2001), Learning By Doing and the Dutch Disease, *European Economic Review 45*, 285-306.

[33] Van der Ploeg, F. (2006), Challenges and Opportunities for Resource Rich Economies, CEPR Discussion Paper 5688.

[34] Van Wijnbergen, S.J.G. (1984), The Dutch Disease: a Disease after All?, *The Economic Journal 94*, 41-55.

[35] World Bank (2006), *Where is the Wealth of Nations? Measuring Capital for the 21st Century*, World Bank, Washington DC.

[36] Zeira, J. (1990), Cost Uncertainty and the Rate of Investment, *Journal of Economic Dynamics and Control 14*, 53-63.

Table 1a Summary statistics

	Obs.	Mean	St. Dev.	Min.	Max.
Household income inequality (Gini index)	1988	41.39	6.70	24.07	59.09
Δ Household income inequality (Gini index)	1988	0.11	1.44	-14.72	11.06
Real GDP per capita (log)	1988	8.02	1.48	4.69	10.73
Political constraints	1988	0.43	0.33	0.00	0.89
Primary schooling (average number of years)	1988	3.94	1.63	0.47	7.70
Commodity export price index	1958	0.33	0.33	0.00	1.87
Unlogged unweighted index (1980=100)	1958	85.70	26.88	17.60	224.48
Δ Commodity export price index	1958	-0.00	0.02	-0.23	0.31
Δ Unlogged unweighted index (1980=100)	1958	-0.79	13.02	-103.10	76.40
Commodity exports to GDP (ratio)	1988	0.08	0.08	0.00	0.35
Δ Non-agricultural commodity export price index	1958	-0.00	0.02	-0.23	0.31
Non-agricultural commodity exports to GDP (ratio)	1988	0.04	0.07	0.00	0.34
Δ Agricultural commodity export price index	1958	-0.00	0.01	-0.07	0.06
Agricultural commodity exports to GDP (ratio)	1988	0.03	0.05	0.00	0.22

Table 1b Sample countries (ISO alpha-3 codes) and their shares of commodity exports in GDP

zmb 0.35	hnd 0.20	cmr 0.15	cri 0.11	sen 0.06	nld 0.05	phl 0.03	prt 0.02	hkg 0.01	bgd 0.01
ven 0.32	tto 0.19	mus 0.15	col 0.11	bdi 0.06	tha 0.04	zaf 0.03	hun 0.02	fra 0.01	aut 0.01
lbr 0.30	isl 0.19	syr 0.15	zwe 0.09	per 0.06	ury 0.04	pol 0.03	brb 0.02	esp 0.01	ita 0.00
cog 0.29	png 0.18	dza 0.15	gtm 0.08	slv 0.06	mex 0.04	irl 0.03	caf 0.02	ind 0.01	kwt 0.00
swz 0.22	jam 0.18	tgo 0.14	dom 0.07	nzl 0.06	gmb 0.04	grc 0.02	fin 0.02	tur 0.01	lso 0.00
sgp 0.21	nic 0.17	irn 0.14	lka 0.07	ken 0.06	uga 0.04	chn 0.02	gbr 0.02	hti 0.01	npl 0.00
ecu 0.21	nor 0.17	bol 0.12	aus 0.07	jor 0.05	rwa 0.04	egy 0.02	cyp 0.02	usa 0.01	ben 0.00
mys 0.21	chl 0.16	fji 0.12	bwa 0.07	can 0.05	dnk 0.04	bra 0.02	swe 0.02	isr 0.01	bel 0.00
mwi 0.20	idn 0.15	gha 0.11	tun 0.06	pan 0.05	arg 0.03	pak 0.02	moz 0.01	kor 0.01	jpn 0.00

Table 2 Estimation results: baseline specifications

	(1)	(2)	(3)	(4)	(5)	(6)
			Estimates of long-run coefficients			
GDP per capita (log)	-2.26	-4.02***	-4.12***	-4.11***	-4.11***	-4.12***
	(1.71)	(1.49)	(1.48)	(1.48)	(1.49)	(1.48)
Political constraints	-0.40	-3.02**	-3.15**	-3.14**	-3.14**	-3.14**
	(3.82)	(1.26)	(1.25)	(1.26)	(1.26)	(1.25)
Primary schooling	-0.15	-1.45*	-1.52*	-1.51*	-1.52*	-1.52*
	(0.93)	(0.82)	(0.86)	(0.82)	(0.82)	(0.82)
Commodity export price index			0.48			
			(6.60)			
			Estimates of short-run coefficients			
Inequality$_{t-1}$	-0.05***	-0.20***	-0.20***	-0.20***	-0.20***	-0.20***
	(0.01)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Δ Inequality$_{t-1}$	-0.18***	-0.14**	-0.14**	-0.14**	-0.14**	-0.14**
	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)
Δ Commodity export price index$_t$			-2.40	-2.45		
			(1.75)	(1.62)		
Δ Non-agricultural export price index$_t$					-3.23**	-3.26**
					(1.53)	(1.54)
Δ Agricultural export price index$_t$					-2.88	
					(6.93)	
Country fixed effects	NO	YES	YES	YES	YES	YES
Regional time dummies	YES	YES	YES	YES	YES	YES
Observations	1988	1988	1958	1958	1958	1958
R-squared (within)	0.17	0.23	0.23	0.23	0.23	0.23

Notes: The dependent variable is the change in inequality in year t. Robust standard errors are clustered by country and are reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% levels.

Table 3 Estimation results: sensitivity analysis

	(1)	(2)	(3)	(4)	(5)	(6)
			Estimates of long-run coefficients			
GDP per capita (log)	1.75	-4.13***	-2.29	-5.22***	-4.08***	-4.09***
	(2.20)	(1.48)	(1.70)	(1.19)	(1.49)	(1.49)
Political constraints	0.07	-3.13**	-0.21	-2.46**	-3.18**	-3.19**
	(2.20)	(1.25)	(3.80)	(0.98)	(1.26)	(1.27)
Primary schooling	1.79**	-1.53*	-0.17	-1.28**	-1.50*	-1.51*
	(0.86)	(0.82)	(0.93)	(0.63)	(0.87)	(0.82)
Commodity export price index					-0.61	
					(6.54)	
Uncertainty					1.33***	1.32***
					(0.42)	(0.45)
			Estimates of short-run coefficients			
Inequality$_{t-1}$	-0.13***	-0.20***	-0.05***	-0.22***	-0.20***	-0.20***
	(0.03)	(0.03)	(0.01)	(0.02)	(0.03)	(0.03)
Δ Inequality$_{t-1}$	-0.15**	-0.14**	-0.19***	-0.07***	-0.14**	-0.14**
	(0.06)	(0.06)	(0.06)	(0.03)	(0.06)	(0.06)
Δ Commodity export price index$_t$					-2.41	-2.34
					(1.74)	(1.62)
Δ Non-agricultural index$_t$	-2.37**		-3.75**	-4.23		
	(1.18)		(1.73)	(2.97)		
Δ Non-agricultural index$_t$-positive		-4.16**				
		(1.77)				
Δ Non-agricultural index$_t$-negative		-2.24				
		(2.93)				
Method	FE	FE	OLS	FE-2SLS	FE	FE
Regional time dummies	NO	YES	YES	YES	YES	YES
Observations	1958	1958	1958	1802	1958	1958
R-squared (within)	0.09	0.23	0.17	0.24	0.23	0.23

Notes: The dependent variable is the change in inequality in year t. Standard errors are reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% levels.

Figure 1 The Edgeworth-Bowley box with dynamically stable skilled and unskilled labor allocations

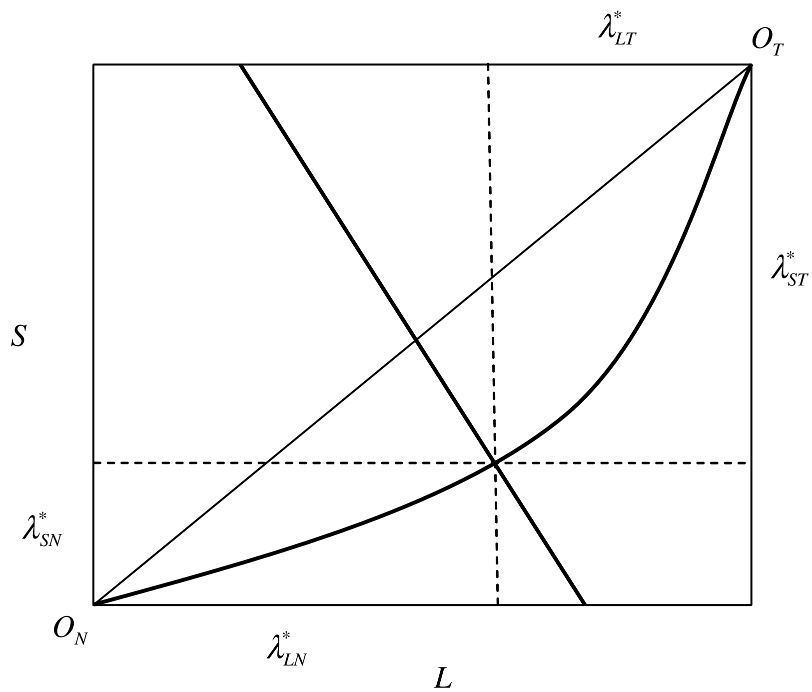

Figure 2 The time path of inequality following a resource boom

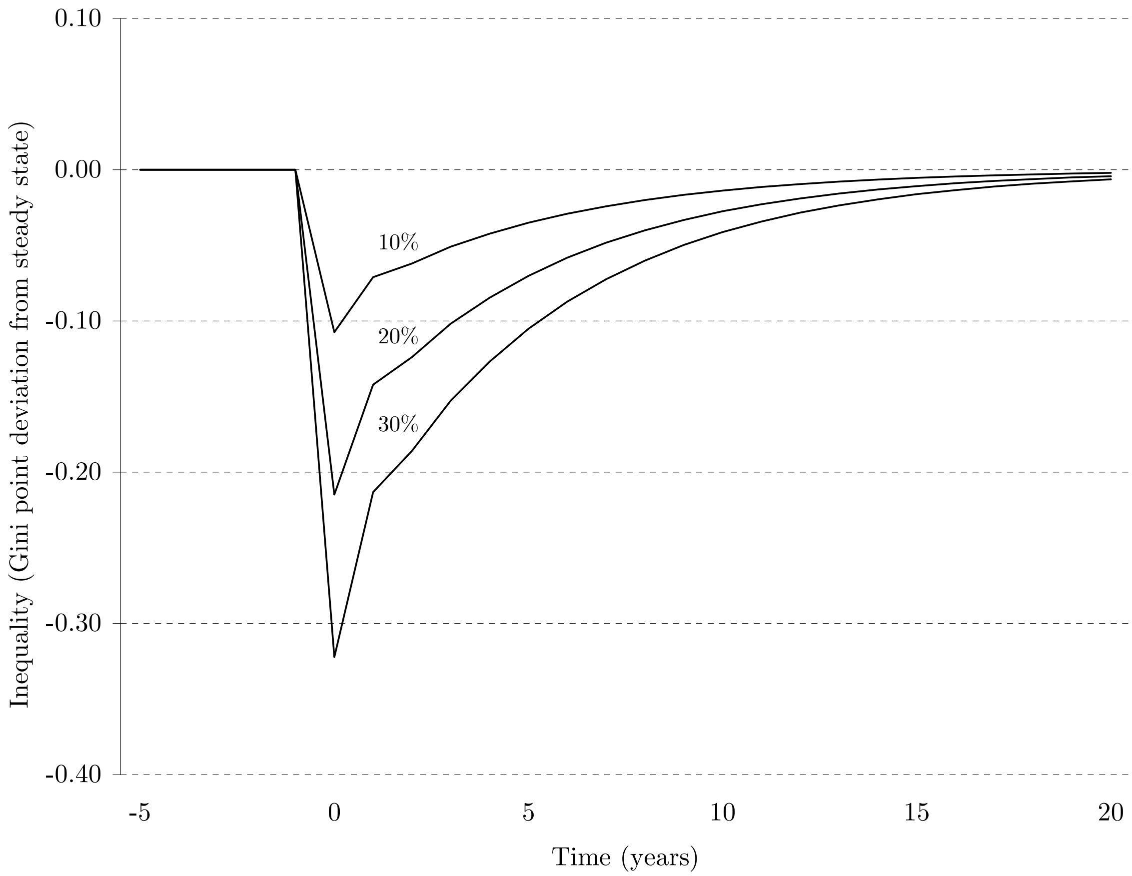

Figure 2 is based on the estimation results in Table 2, column (6). It shows impulse response functions of inequality for a two standard deviations (33% points) shock to the growth rate of non-agricultural export prices in period 0, for countries with shares of non-agricultural exports in GDP of 10%, 20%, or 30%.