

# Spend, Baby, Spend: Windfalls, Specialization and Misallocation<sup>1</sup>

Radoslaw (Radek) Stefanski  
University of St Andrews and OxCarre

October 12, 2014

## Abstract

Resource-rich countries tend to employ a much larger proportion of workers in the government sector than resource-poor countries. I construct a model with a productive government sector and examine optimal government employment in resource-rich countries. In the calibrated best-case-scenario model, predicted optimal government employment is nearly 10 times smaller than in the data. This implicit misallocation of resources has a large impact on welfare and aggregate productivity. Using the calibrated model I find that a ten percentage point increase in resource windfalls is associated with a 1.94% percent lower aggregate productivity and a 1.22% lower welfare arising from government misallocation.

---

<sup>1</sup> This paper has been prepared for the XVIII Central Bank of Chile Annual conference. Thanks to Antonio Mele for useful discussion and comments. All errors are my own. Email: rls7@st-andrews.ac.uk.

# 1 Introduction

In this paper I investigate the impact of structural transformation in an open economy on sectoral and aggregate productivity with a particular focus on the role of government. Structural transformation is a reallocation of labor across sectors. Whilst there are potentially many sources of structural transformation,<sup>2</sup> I focus on labor reallocation induced by a windfall of revenue. Furthermore, I concentrate only on windfall revenue arising from the export of natural resources (fuels, ores and metals), although the entire analysis is applicable to other types of windfalls such as - for example - foreign aid, remittances, EU structural funds or war reparations.

The exact focus here is the size of public sector employment in resource-rich countries. Governments provide largely non-traded services (such as law enforcement, defense, infrastructure, arbitration etc.) and thus we can expect the standard, “Dutch-Disease” mechanism pushing workers towards non-traded sectors in resource-rich countries to hold: higher windfalls of revenue should increase demand for both traded and non-traded goods, but since the supply of non-traded goods can only be provided locally, more workers need to shift to non-traded sectors - including the government sector - in order to satiate the higher demand for non-traded goods in resource-rich countries. As such, I am interested in how the size of public employment *should* vary optimally between resource-rich and resource poor countries, whether the extent of government employment observed in resource-rich countries is efficient and - if not - what the productivity and welfare costs are of this misallocation.

I do two things in this paper. First, using a panel of macro cross-country data I demonstrate that the share of public sector employment is greater in resource-rich countries than in resource poor countries - even controlling for the size of other non-traded sectors. Second, I construct and calibrate a small, open economy model with two production sectors and a government sector in which (optimally) higher government employment shares emerge as a consequence of windfall-induced labor reallocation. I then use a model to compare the optimal and observed size of government in order to obtain an estimate of the extent of government misallocation and the impact it has on welfare and productivity.

Importantly, the paper builds on earlier work by Kuralbayeva and Stefanski (2013). In that paper we did two things. First, we showed that resource-rich regions tend to have a) *small* but *relatively productive* manufacturing sectors and b) *large* but *relatively unproductive* non-manufacturing sectors. Whilst this difference in sectoral size (or Dutch-Disease effect) was well known and in line with theoretical predictions,<sup>3</sup> the productivity facts were novel and we showed that standard models were ill-equipped to replicate them. Second, we constructed and

---

<sup>2</sup> Gollin et al. (2002), Duarte and Restuccia (2010), Rogerson (2008), Dekle and Vandenbroucke (2011) and Yi and Zhang (2010), for instance, focus on labor reallocation induced by non-homotheticities in agriculture.

<sup>3</sup> See for instance, Corden and Neary (1982), Matsuyama (1992) or Michaels (2011) for theoretical and empirical treatments of this so-called Dutch Disease.

calibrated a small, open economy model with two sectors in which observed differences in sectoral productivity emerged endogenously as a consequence of windfall-induced labor reallocation and subsequent worker specialization. Since, in the current paper, I am interested in studying the impact of windfall-induced changes in government size on sectoral and aggregate productivity, it is crucial to correctly capture the windfall induced changes in sectoral productivity that are *not* driven by changes in the size of the government sector. As such, in this paper I adapt the framework of Kuralbayeva and Stefanski (2013) which does well in reproducing the pertinent facts relating to both sectoral size and sectoral productivity in resource-rich countries in the absence of government.

More specifically, in my model, I assume that manufacturing consumption goods are traded whilst non-manufacturing consumption goods are non-traded and that agents have heterogeneous skills at performing different tasks in both consumption-good sectors. In addition, I assume the existence of government sector whose role it is to provide basic public services such as an institutional framework, law-enforcement, judiciary, defense, infrastructure etc. to the two consumption good sectors. The government sector is modeled as having a positive (external) effect on the productivity of both consumption sectors - however, government employees will have to be paid through a tax levied on all workers. I will also assume that government services cannot be imported from abroad. A region with higher windfall revenues will demand more of both consumption goods and of government services than a region without windfalls. Whilst the region's higher demand for manufacturing consumption goods can be satiated by imports from abroad, more workers need to be employed in non-manufacturing sectors (including the government sector) to meet the higher demand for locally produced non-manufacturing consumption goods and government services. This generates a reallocation of labor from manufacturing to the non-manufacturing sectors and results in a process of self-selection. Workers who choose to remain in manufacturing despite a windfall are those who are most skilled at manufacturing sector tasks, which leads to a more specialized and hence a more productive manufacturing sector. Workers who re-allocate to non-manufacturing do so only in response to the higher demand generated by the windfall and will be less skilled at non-manufacturing sector tasks than workers already employed in that sector. This can lead to a more de-specialized and hence less productive non-manufacturing sector.<sup>4</sup> Windfalls thus induce labor reallocation which in turn can generate asymmetric changes in sectoral productivity and an increase in the size of government.

I calibrate the model and show that the exogenous variation in endowments of natural resources does remarkably well in explaining the differences in sectoral employment structure and the large, asymmetric differences in sectoral productivity observed across countries. The model

---

<sup>4</sup> Although the extent of this de-specialization can be tempered by the higher productivity resulting from a bigger government - the exact pattern will depend on the particular calibration

also does well in explaining differences in non-manufacturing prices in the data. However, the optimal increase in government employment in resource-rich countries predicted by the model is significantly smaller than the employment observed in the data. resource-rich countries seem to employ far more workers in government than the above model would suggest is optimal. In order to calculate the cost of this apparent misallocation, I feed in observed government employment levels into my model, and examine the subsequent changes in labor reallocation across manufacturing and non-manufacturing sectors and the resultant differences in productivity. I find that a ten percentage point increase in resource windfalls is associated with a 1.94% percent lower aggregate productivity and a 1.22% lower welfare arising from government misallocation. In short, resource-rich countries tend to have governments that are too big, and this can have a relatively large impact on both productivity and welfare.

The above idea of a negative relationship between natural resources and economic outcomes ties into the so-called “resource curse” literature - see for example Neary (1978), van der Ploeg (2010), Robinson et al. (2006), Collier and Goderis (2009), Collier and Hoeffler (2005) etc. Whilst the conclusions of that literature are not definitive, there is strong evidence to suggest that resource windfalls can generate various negative economic effects especially in the presence of bad governance and poor institutions. In particular, in that literature, negative economic outcomes are often a consequence of a corrupt political process associated with higher resource wealth. In short, those papers tend to argue that resource-rich countries offer more opportunities for graft which introduces a drag on the economy. The approach taken in this paper is different and intentionally complimentary. In the model, I take the most charitable view of government possible. In particular, I assume that the government sector is a crucial input in production and that there is no corruption, no directly wasted resources, no electioneering, no graft and no costly power struggles. Governance in resource-rich countries is also assumed to be potentially just as effective as in resource poor countries. Thus, I do my best to give governments in resource-rich countries the benefit of the doubt and as such my model aims to generate the largest possible optimal increase of government employment in response to windfalls. In my setup, the only way that government can be inefficient is if it employs too many or too few workers relative to what is predicted optimal by the model. Importantly, however, I do not take a stand on *why* governments are the size that they are and instead - in my baseline experiment - I simply take public sector employment from the data and analyze the implicit misallocation costs of governments that are too big or too small.

Like Kuralbayeva and Stefanski (2013), the self-selection aspect of this work is in the spirit of Lagakos and Waugh (2012), Roy (1951) and Lucas (1978) and is closely linked to a similar discussion in the development literature. Poorer countries tend to have a larger fraction of their labor force employed in agriculture, due to subsistence requirements. Caselli (2005) and

Restuccia et al. (2008) also show that productivity differences in agriculture between rich and poor countries are significantly greater than aggregate productivity differences. Lagakos and Waugh (2012) argue that this fact stems from the specialization that takes place in the smaller agricultural sectors in rich countries. They formalize and test their idea in the framework of a Roy (1951) model of self-selection. The outcomes of the above models however are efficient and do not consider the impact of a misallocation stemming from suboptimal government size. Furthermore, Lagakos and Waugh (2012) rely on non-homothetic preferences and an exogenous variation in aggregate productivity to generate a shift of workers across sectors. The current model has homothetic preferences and instead emphasizes the role of exogenous resource wind-falls and the existence of a non-traded sector as the channel driving labor reallocation. Thus I avoid what Lagakos and Waugh (2012) call the “key challenge” of their setup which is the requirement of large, exogenous productivity differences to drive workers across sectors.

Section 2 introduces the data used in this study and establishes the productivity and employment facts. Section 3 introduces a general version of the model whilst section 4 considers the role of heterogeneity and government in a simple version of the model. Sections 5 and 6 present the solution and calibration of the general model, section 7 presents the results, whilst section 8 delves into the scope of the government misallocation and its impact on productivity and welfare. Finally, I conclude in section 9.

## 2 Data and Facts

In this section I briefly review the data and facts pertaining to manufacturing and (non-resource) non-manufacturing employment shares and productivity constructed in Kuralbayeva and Stefanski (2013). I also examine the data and facts pertaining to employment in the government sector. In particular, I show that 1) resource-rich regions have small and relatively productive manufacturing sectors, 2) large and relatively unproductive non-manufacturing sectors and 3) a greater proportion of workers employed in the government sector.

Throughout, I follow Kuralbayeva and Stefanski (2013) and divided economies into mining and utilities ( $MU$ ), manufacturing ( $M$ ) and non-resource non-manufacturing ( $NM$ ) sectors:<sup>5</sup>

$$\text{Total Economy} = \overbrace{\underbrace{A + C + S + G}_{\text{Non Res. Non-Mfg.}} + \underbrace{M}_{\text{Mfg.}}}^{\text{Non-Resource Economy}} + \underbrace{MU}_{\text{Mining and Utilities}}. \quad (1)$$

As in Kuralbayeva and Stefanski (2013), I focus only on the productivity and employment

---

<sup>5</sup> The lowest level of aggregation available for all data is the one sector ISIC classification. NM here is defined as the sum of agriculture (A), construction (C) and (private) services (S) and Government (G).

structure of the non-resource economy.<sup>6</sup> Diverging from Kuralbayeva and Stefanski (2013) however, I will also consider the proportion of non-manufacturing workers employed in the public sector. Notice, however, that I will not say anything about productivity in the government sector. Constructing sectoral productivity measures is challenging and the assumptions needed to calculate government-specific productivity would be heroic to say the least. In what follows I give a brief overview of the data.

**Data** In Kuralbayeva and Stefanski (2013) we considered different measures of productivity for the manufacturing and non-manufacturing sector. Here, I will only consider the most comprehensive measure of productivity from that paper,  $D_s$ , obtained as a residual from the following production function:

$$Y_s = D_s(K_s)^{\alpha_s}(h_sL_s)^{1-\alpha_s} \quad (2)$$

where  $Y_s$  is sector  $s$ 's value-added,  $L_s$  is sectoral employment,  $K_s$  is sectoral physical capital and  $h_s$  is average sectoral human capital, so that  $h_sL_s$  is the 'quality adjusted' workforce.<sup>7</sup> Constant price sectoral value-added data comes from the UN, and is adjusted to control for cross-country sectoral price level differences using the World Bank's 2005 International Comparison Program (ICP) price data. Employment data comes from the ILO and physical capital is constructed using the perpetual inventory method from the PWT. I follow Caselli (2005) in constructing aggregate human capital from the Barro and Lee (2010) education data set and in constructing sectoral physical capital. Finally, due to lack of data, I assume that the ratio of human capital between any two sectors is constant across countries and time and equal to the corresponding ratio in the US and that labor shares in the last two measures of productivity,  $1 - \alpha_s$ , are identical across countries, constant over time and equal to OECD averages. For construction details, see the Appendix of Kuralbayeva and Stefanski (2013).

Next, I obtain public sector employment data from the ILO which "covers all employment of (the) general government sector as defined in System of National Accounts 1993 plus employment of publicly owned enterprises and companies, resident and operating at central, state (or regional) and local levels of government. It covers all persons employed directly by those institutions, without regard for the particular type of employment contract."<sup>8</sup>

I follow Sachs and Warner (2001) and Kuralbayeva and Stefanski (2013) in defining natural resource "wealth" as the ratio of exports of natural resources (fuels, ores and metals) to GDP

<sup>6</sup> Thus, when we refer to aggregate productivity or sectoral employment share, we always mean aggregate productivity of the *non-resource* economy or sectoral employment relative to *non-resource* employment.

<sup>7</sup> I also refer to  $D$  as the corresponding measures of aggregate (non-resource) productivity.

<sup>8</sup> A limited subset of the public employment data is provided at the ISIC one sector level and in that (very limited) subset, public employment is overwhelmingly in the non-manufacturing sector. As such in the baseline experiment of this paper - in order to maintain as large a sample of data as possible - I shall assume that all government employment belongs entirely to the non-manufacturing sector.

	E Res./	Output/	Emp. Share		Emp. Share		TFP ( $D_s$ )	
	GDP	worker	NM	M	G	NG	NM	M
10 <sup>th</sup> %-ile	0.17	30,716	0.86	0.14	0.25	0.75	0.94	1.37
90 <sup>th</sup> %-ile	0.00	17,756	0.81	0.19	0.17	0.83	0.98	1.11
10 <sup>th</sup> /90 <sup>th</sup>			1.06	0.73	1.48	0.90	0.96	1.24

Table 1: Resource export share, output per worker, composition of employment (manufacturing vs. non-manufacturing and public vs. non-public) and a measure of sectoral productivity (relative to aggregate productivity) for top and bottom 10% of natural resource exporters. Data for 33 countries, for 1980-2007. (Source: see Section 2)

using WDI data.

In my baseline sample, like in Kuralbayeva and Stefanski (2013), I consider a panel of the 120 richest countries for the years 1980-2007 period.<sup>9</sup> I keep all country-date points for which I have all necessary data and those that do not deviate significantly across different data sources. This leaves me with a total of 33 countries in my sample. On average, there are 10 observations for each country, 22 observations for each year and a total of 340 data points. Notice that until 1995, the data for public employment is only available once every five years and there are very few observations from 1980 and 1985.

**Summary of Facts** Table 1 shows summary results by comparing the largest 10 percent of natural resource exporters with the smallest 10 percent. The table reproduces the results (pertaining to sectoral size and productivity) found in Kuralbayeva and Stefanski (2013) for the current sample of data and adds the new finding pertaining to the size of government employment in resource-rich countries. The table shows the decomposition of employment according to manufacturing/non-manufacturing and public/non-public sectors. Furthermore it also shows the sectoral productivity (in manufacturing and non-manufacturing) normalized by aggregate productivity of each group. From the table observe that resource-rich countries: 1) employ proportionally 27% less workers in manufacturing and 6% more workers in (non-resource) non-manufacturing than resource-poor countries and 2) that they are 24% more productive in manufacturing and 4% less productive in non-manufacturing (relative to aggregate productivity) than resource-poor countries. Finally, also notice that resource-rich countries: 3) employ 48% more workers in the public sector and 10% less workers in the non-public sector than resource-poor countries.

<sup>9</sup> I focus on richer countries for three reasons: First, I am examining more disaggregate data than is standard so data quality in poorer countries is a serious concern. Second, I feel that the mechanism of specialization described later may play a more prominent role in richer countries. Finally, focusing on richer countries may avoid the worst of unobserved cross-country heterogeneity. Since this procedure may in principle result in unobserved selection bias, I have also experimented with a complete sample and the results are independent of this cutoff.

	(1)	(2)	(3)	(4)
	M. Emp.	$\log(D_m)$	$\log(D_s)$	$\log(p_s)$
log(NRE)	-0.014*** (0.002)	0.068*** (0.014)	-0.012*** (0.002)	0.048*** (0.011)
logLprod	0.650*** (0.127)			
sqlogLprod	-0.031*** (0.006)			
log(D)		1.458*** (0.078)	0.888*** (0.011)	0.838*** (0.067)
$\tau_e$				-0.355*** (0.132)
Time FE	yes	yes	yes	yes
Obs.	340	340	340	340
$R^2$	0.256	0.567	0.953	0.480

Standard errors in parentheses  
\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Table 2: Baseline results from Kuralbayeva and Stefanski (2013) in current sample of data.

As in Kuralbayeva and Stefanski (2013), I stress that the productivity results refer to *relative* and *not absolute* productivity. So, for example, looking at the column labeled  $D_m$  in Table 1, the average productivity of manufacturing in the top 10% of resource exporters is 37% higher than the average aggregate productivity of those same countries, whereas in the bottom 10% of exporters the average manufacturing productivity is only 11% higher than the average aggregate productivity in that group of countries. Countries that have low aggregate (or sector neutral) levels of productivity will have low *absolute* levels of productivity in all sectors irrespective of the size of their resource endowments but may still have high productivity in manufacturing *relative* to their aggregate productivity.

**Earlier Results** In this section in Table 2, I briefly reproduce the baseline regressions of Kuralbayeva and Stefanski (2013) for the current sample of data. For robustness with respect to these regressions and further discussion see that paper. Column (1) shows the regression of manufacturing employment share on the log of the windfall measure controlling for changes in output per worker (and output per worker squared) as well as controlling for time-fixed effects.<sup>10</sup>

Resource-rich countries employ less workers in the manufacturing sector and (implicitly) more

<sup>10</sup> Since employment share in manufacturing is simply one minus the employment share in non-manufacturing, the regressions for non-manufacturing employment are the same with opposite signs on coefficients. As in Kuralbayeva and Stefanski (2013), I take a log transformation of resource windfalls since the data is concentrated near zero. This ensures that the transformed empirical distribution is closer to normal. Importantly this transformation does not drive the results. Finally, I control for output-per-worker and output-per-worker squared since it is

workers in the non-manufacturing sector: a doubling of resource windfalls is associated with a 1.4 percentage point decline in the manufacturing employment share. These results are statistically significant at the one percent level.

Columns (2) and (3) of Table 2 show how (the log of) manufacturing and non-manufacturing productivity varies with (the log of) resource windfalls and aggregate productivity. Higher aggregate (or sector neutral) productivity is unsurprisingly associated with higher sectoral productivity. However, controlling for differences in aggregate productivity, resource-rich countries tend to be more productive in manufacturing and less productive in non-manufacturing than resource-poor countries. These results are significant at the one percent level and are robust to other measures of productivity as discussed in Kuralbayeva and Stefanski (2013). A doubling of natural resource windfalls is associated with a 1.2% lower non-manufacturing productivity and a 6.8% higher manufacturing productivity.<sup>11</sup>

An important fact that will be later, is the positive impact of windfalls on the non-manufacturing price. In Kuralbayeva and Stefanski (2013) we constructed a panel of sectoral price level data by combining ICP cross-country sectoral price levels with sectoral price indices from the UN. Column (4) reproduces the baseline price regression from Kuralbayeva and Stefanski (2013). In particular it shows the regression of the log of relative non-manufacturing price levels (with respect to manufacturing price levels) on the (log) measure of resource windfalls, (log) aggregate productivity, energy subsidies from WEO<sup>12</sup> and time-fixed effects.<sup>13</sup> I find that a doubling of natural resource windfalls is associated with a 4.8% increase in the price of non-traded goods and these results are significant at the one percent level.

**Public Sector Employment Results** Next, I present the novel empirical results of this paper. Table 3 shows the regressions relating the size of the government sector employment with resource windfalls. In particular, column (1) shows the regression of government employment share on the log of my windfall measure. Resource-rich countries employ more workers in the public sector and (implicitly) less workers in the non-public sector. These results are statistically

---

a well established fact that manufacturing follows a hump shape with income. This in no way drives my results. For details and robustness tests, see Kuralbayeva and Stefanski (2013).

<sup>11</sup> I emphasize that these results refer to *relative* and *not absolute* productivity. Countries that have low aggregate (or sector neutral) levels of productivity will have low *absolute* levels of productivity in all sectors irrespective of the size of their resource endowments but may still have high productivity in manufacturing *relative* to aggregate productivity.

<sup>12</sup> Subsidy data is an average of 2008-2010 data. We assume that these subsidies are country specific and fixed over the 1980-2007 period.

<sup>13</sup> Notice that we included aggregate productivity to control for the so-called Penn effect - the observation that richer countries have higher non-traded goods prices than poorer countries. Furthermore, as was discussed in Kuralbayeva and Stefanski (2013) a potential issue with the ICP price data is that they reflect consumer rather than producer prices - which are the focus of the later model. This may be particularly important in resource-rich economies, where consumer subsidies are prevalent. We control for energy subsidies as an indirect attempt at controlling for the overall level of subsidies in a country's economy

	(1)	(2)	(3)	(4)	(5)
	G. Emp.				
log(NRE)	0.014*** (0.004)	0.017*** (0.004)	0.011*** (0.004)	0.013*** (0.004)	0.021*** (0.004)
logLprod			0.041*** (0.008)	0.045*** (0.008)	0.045*** (0.008)
acs					-0.595*** (0.103)
Time FE	no	yes	no	yes	yes
Obs.	340	340	340	340	340
$R^2$	0.045	0.061	0.115	0.139	0.220

Standard errors in parentheses  
\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Table 3: Changes in government employment share and resource wealth.

significant at the one percent level. Column (2), controls for time fixed effects, whilst column (3) controls for changes in output per worker which may help reduce unobserved cross-country heterogeneity. Column (4) adds time-fixed effects to the regressions in column (3). In all three cases, the results remain largely unchanged. Finally, column (4) adds employment shares of the non-manufacturing sector. The results of this last regression tells us that - even controlling for the size of other non-manufacturing sectors - resource-rich countries tend to have a larger government sector. Taking column (2) as the baseline result, I find that a doubling of natural resource windfalls is associated with a 1.7% higher public sector employment share and these results are significant at the one percent level.

Finally, notice that whilst in Kuralbayeva and Stefanski (2013) we controlled for time and country-fixed effects, in the above regressions I include only time-fixed effects. There are two reasons for this. First, I have a far more limited sample of data and so there is not enough variation over time in the sample. Most of the variation over time in windfalls comes from variation in price which tends to be common across countries. Since much of the price variation in natural resources took place in the 1980's and much of our public employment data is missing in that period, there is very little temporal variation in the remaining data.<sup>14</sup> Second, and perhaps more importantly, the focus of this paper will be government employment. This type of work is often characterized by tenure or unionization and is thus often quite unresponsive to shocks over time - at least in the short run. As such, to examine the persistent effects of resource endowments, it makes more sense to look at cross-country differences which can be interpreted

<sup>14</sup> A rule of thumb here is to regress the independent variable -  $\log(NRE)$  - on country fixed effects. If the value of  $1/(1 - R^2)$  from the resulting regression is less than ten, the rule of thumb suggests that there is enough variation in the data to include that variable. In our case  $1/(1 - R^2) = 11$  thus suggesting there is too little variation to include country fixed effects.

as long run effects.

### 3 Model

In this section I introduce a small, open, multi-sector economy with heterogenous agents that can account for the observed facts in productivity and employment. The model closely follows Kuralbayeva and Stefanski (2013), but introduces a role for government. There are three final goods in the economy: manufacturing goods ( $m$ ), private non-manufacturing goods which - for brevity - I will call services ( $s$ ) and a windfall good which, also for brevity, I will refer to as oil but could equally well be any other natural resource or alternative source of windfall revenue. I assume that manufacturing and oil are traded internationally, whilst services are assumed to be non-traded. Oil is assumed to be an endowment good that is not used locally but only exported abroad (and thus serves as a windfall of income), whilst manufacturing and services can be produced locally using labor but no oil. I also assume the existence of a government sector (the public non-manufacturing sector) which provides the manufacturing, service and oil sectors with inputs such as institutional frameworks, transportation, rule of law, arbitration etc. that are productivity enhancing, but are external to firms (and workers). Thus, whilst workers can be employed in the government sector, the sector produces no final goods directly, but rather provide an input that looks like a higher level of productivity to other sectors of the economy. Finally, I assume that the external benefits produced by government cannot be imported from abroad.

**Households** Suppose there is a measure one of agents, indexed by  $i$ . Preferences are given by:

$$U(c_s^i, c_m^i) \equiv \left( (c_s^i)^{\frac{\sigma-1}{\sigma}} + \nu(c_m^i)^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}, \quad (3)$$

Each agent in the economy is endowed with a unit of time and assumed to have a vector of innate sector specific skills or talents,  $\{z_s^i, z_m^i\}$ , representing the efficiency of that unit of time in the service sector ( $s$ ) and the manufacturing ( $m$ ) sector. Endowments of skills  $\{z_s^i, z_m^i\}$  are exogenous and are assumed to be randomly drawn from a distribution common to the whole population  $N(z_s, z_m)$ . Since skills are assumed to be perfectly observable, agents earn a wage income,  $w^i$ . The agent is also endowed with a resource tree that provides a stream of  $O$  units of oil each period. Oil is not directly used by the agent but is exported and provides windfall revenues. Finally, each agent also potentially faces a lump-sum tax,  $T$ , paid to government. The budget constraint of the agent is thus given by:

$$p_s c_s^i + c_m^i \leq w^i + G_m(L_g)p_o O - T, \quad (4)$$

where,  $p_s$  is the relative price of service sector goods and  $p_o$  is the relative price of oil determined on international markets. Traded manufacturing goods are taken as numeraire. Finally, in the above,  $0 \leq G_m(L_g) \leq 1$  is a function capturing the external productivity benefits of government to the export of oil. These are assumed to be positively dependent on the employment size of the government sector,  $L_g$ . I described this function in more detail in following paragraphs.

**Production** I assume a competitive market in all final good sectors so that each worker gets paid his marginal product. The output of worker  $i$  in sector  $k = s, m$  is given by  $Y_k^i = AG_k z_k^i$ , where  $A$  is aggregate (potentially sector specific) efficiency,  $z_k^i$  is the worker's sector specific productivity and  $0 \leq G_k(L_g) \leq 1$  is the impact of government on productivity that is external to workers and firms but depends positively on the size of government employment,  $L_g$ , in a manner described in the following paragraph. Aggregate output in sector  $k = s, m$  is given by:

$$Y_k \equiv \int_{i \in \Omega^k} Y_k^i di = AG_k(L_g) \tilde{L}_k, \quad (5)$$

where  $\Omega^k$  is the set of agents electing to work in sector  $k$ ,  $L_k \equiv \int_{i \in \Omega^k} di$  is the number of workers in private enterprises in sector  $k$  and  $\tilde{L}_k \equiv \int_{i \in \Omega^k} z_k^i di$  represents the total effective labor units (privately) employed in sector  $k$ . Finally, notice that for simplicity, I assume that  $G_m(L_g)$  is common to both the oil sector and the manufacturing production sector.

**Trade** It is assumed that manufacturing goods and oil are traded whilst service sector goods are not traded. In order to close the model, I assume a period-by-period balanced budget constraint given by:

$$m - G_m(L_g)p_o O = 0, \quad (6)$$

where,  $m$  is the value of imported traded goods (recall that traded goods are assumed to be the numeraire) and  $G_m(L_g)$  is the impact of government on how effective imports are, capturing the idea of a type of iceberg transport cost. As mentioned above, for simplicity I have assumed that the government contribution to the productivity of exporting (or producing) oil is the same as the corresponding term in the manufacturing sector. Finally, in the above setup, all oil endowments are exported in exchange for manufacturing imports. A country with no oil (i.e.  $p_o O = 0$ ) is thus assumed to be closed to trade.

**Government** The government employs  $L_g$  workers to provide public goods and services such as infrastructure, a justice system, law and order etc. that enhance sectoral productivity of the consumption sector,  $G_k(L_g)$ , according to the following production function:

$$G_k(L_g) = 1 - \frac{\psi^k}{\psi^k + L_g}, \quad (7)$$

where  $\psi^k \geq 0$  is a sector specific constant capturing the importance of government services to production in a particular sector. When  $\psi^k > 0$ ,  $G'_k(L_g) > 0$  - that is more government employees contribute more - ceteris paribus - to the output of a sector. Zero employment in the government sector implies  $G_k(0) = 0$  and hence zero output in sector  $k$ . Consequently, with  $\psi^k > 0$ , a positive employment in government is necessary for production to take place. If  $\psi^k = 0$ , then  $G_k(L_g) = 1$  and the model collapses to the no-government world of Kuralbayeva and Stefanski (2013). I let  $\Omega^G$  be the set of workers employed in the government sector whilst the number of workers employed in government is given by  $L_g \equiv \int_{i \in \Omega^G} di$ . Finally, for simplicity, and to capture the inherent equity of government employment, I assume that government pays each employee the same wage,  $w^g$ . Alternatively, we can think of this as a technological constraint either on the ability of government to observe worker specific skills or on the fact that production in the government sector requires a constant level of skill. As such, the government's budget constraint, which is assumed to be balanced period by period, is given by:

$$T = w_g L_g. \quad (8)$$

Thus, the government levies a per period lump-sum tax on each worker to pay for the wages of all it's employees.

**Market Clearing** Defining  $\Omega = \Omega^m \cup \Omega^s \cup \Omega^G$ , the market clearing conditions for manufacturing, services and employment are given by:

$$\int_{i \in \Omega} c_m^i di = Y_m + m \text{ and } \int_{i \in \Omega} c_s^i di = Y_s \text{ and } L_m + L_s + L_g = 1. \quad (9)$$

**Competitive Equilibrium** For each price of oil,  $p_o$ , every endowment level of oil  $O$ , and for a given size of government  $L_g$ , equilibrium in the above economy consists of a relative price of service goods,  $p_s$ , agent-specific wages  $w^i$  and allocations for all agents, firms and government so that labor and output markets clear, and trade as well as the government budget constraint remains balanced, period by period.

**Solution** Each manufacturing and service sector firm chooses a non-negative quantity of labor to hire. Due to perfect competition, firms offer the following wage schedule to consumer  $i$ :

$$w_m^i = AG_m(L_g)z_m^i \text{ and } w_s^i = p_s AG_s(L_g)z_s^i, \quad (10)$$

in manufacturing and service sectors respectively. Consumer  $i$  who decides to work in a non-governmental sector, choose employment in the sector that provides a higher wage given his particular talent vector. The wage offer for each worker in non-governmental work is thus given by,  $w_{ng}^i = \max\{w_s^i, w_m^i\} = \max\{p_s AG_s(L_g)z_s^i, AG_m(L_g)z_m^i\}$  which gives rise to the following

simple cut-off rule: a worker  $i$  employed in non-government work, will choose to work in services if and only if

$$p_s > \frac{G_m(L_g)z_m^i}{G_s(L_g)z_s^i}. \quad (11)$$

Finally, given a worker's wage offer in the private sector, a worker will choose to work in government if he receives a higher wage there. Consequently, the wage of each worker is given by,  $w^i = \max\{w_{ng}^i, w_g\}$ .

Agents take as given prices as well as the wage offers arising from the firm and government problems (and hence the above decision rules). Having picked their specialization, they then proceed to maximize (3) subject to (4), which results in the following demands of each agent:

$$c_s^i = \frac{(w^i + G_m(L_g)p_oO - T)}{p_s + \nu^\sigma p_s^\sigma} \text{ and } c_m^i = \frac{\nu^\sigma p_s^\sigma (w^i + G_m(L_g)p_oO - T)}{p_s + \nu^\sigma p_s^\sigma}. \quad (12)$$

Using the goods market clearing conditions in equation (9) and the demands of each agent from equations (12), I can show that:

$$\nu^\sigma p_s^\sigma Y_s = Y_m + G_m(L_g)p_oO \quad (13)$$

Substituting (5) into (13), provides an implicit expression for  $p_s$  as a function of the value of oil endowment,  $p_oO$  and the level of government employment.<sup>15</sup>

**Optimal Government** So far, I have taken the size of government employment as fixed - and this shall serve as one of the experiments later in the paper. However, it is also possible to suppose that the government is benevolent and wishes to maximize the utility of workers. In particular, a benevolent government will take the demand functions of agents (derived above in equations (12)) as given and solves a Ramsey-type problem for the optimal size of government employment,  $L_g^{opt}$ , by maximizing the expected utility of workers:

$$\max_{0 \leq L_g \leq 1} \mathbb{E}(U(c_s^i(L_g), c_m^i(L_g))), \quad (14)$$

where  $\mathbb{E}(U(c_s^i, c_m^i)) = \int_0^1 U(c_s^i(L_g), c_m^i(L_g)) di$ .

<sup>15</sup> Notice that I have assumed that windfall income are distributed evenly across agents. This assumption plays no role in our results since equation (13) (and hence the equilibrium price and cutoff condition) holds regardless of how windfalls are distributed.

## 4 Heterogenous Workers

### 4.1 A Simple Example

To illustrate the impact of worker heterogeneity and government on sectoral productivity, I begin with a simple example.<sup>16</sup> Suppose the skill distribution  $N$  is degenerate and given by  $\{z_s^i, z_m^i\} = \{e^i, e^{1-i}\}$  for each worker  $i \in [0, 1]$ . Furthermore, assume Cobb-Douglas utility ( $\sigma = 1$ ), equal utility weights ( $\nu = 1$ ), normalize  $A$  to unity and suppose that  $\psi_m = \psi_s > 0$  so that  $G \equiv G_s = G_m$ . Agent  $i$  receives wage offers  $w_s^i = p_s G z_s^i$  in services,  $w_m^i = G z_m^i$  in manufacturing and  $w_g$  in the government sector and will choose to work in the sector that pays most. This gives rise to two cutoff agents,  $\bar{i}_m$  and  $\bar{i}_g$  who are respectively indifferent between manufacturing and government sectors, so that  $w_g = w_m^{\bar{i}_m}$ , as well as between service and government sectors, so that  $w_g = w_s^{\bar{i}_g}$ . Suppose that government hires  $L_g$  workers. To do so it will have to offer a wage large enough so that  $L_g = \bar{i}_g - \bar{i}_m$ . Using these relationships I can calculate these two cutoffs as a function of price and and the size of the government employment so that  $\bar{i}_m(p_s, L_g) = \frac{1 - \log p_s - L_g}{2}$  and  $\bar{i}_g(p_s, L_g) = \frac{1 - \log p_s + L_g}{2}$ . I illustrate the problem of the worker in Figure 1(a) which plots the wage offers in each sector and the cutoffs,  $\bar{i}_k(p_s, L_g)$  for  $k = m, g$ . Agents to the left of  $\bar{i}_m(p_s, L_g)$  are relatively more skilled in manufacturing sector tasks and hence have higher wage offers than in services or government and hence choose to work in the manufacturing sector. Agents to the right of  $\bar{i}_s(p_s, L_g)$  are relatively more skilled in service sector tasks and hence have higher wage offers and choose to work in services. Agents in between the cutoffs, will have a comparative advantage in government work and will hence choose to work in the government sector.

The cutoff values are dependent on the price of service goods and the size of government employment. For the moment, suppose that government adjusts its wage to maintain a constant level of employment and consider the impact of a higher oil windfall. A windfall will influence the price of services and hence the distribution of workers across sectors. A windfall of revenue generates a greater demand for both types of consumption goods. To satiate the higher demand for non-traded service sector goods, more workers are needed in the service sector. New workers however will choose to work in services only if the service wages rise - which in turn can only happen if the service sector price increases. More formally, I can write output in each sector as a function of the respective cutoff (and hence the price):  $Y_s(p_s; L_g) = G_s(L_g) \left( e - e^{\bar{i}_s(p_s; L_g)} \right)$  and  $Y_m(p_s) = G_m(L_g) \left( e - e^{1 - \bar{i}_m(p_s; L_g)} \right)$ . Using these equations as well as the relationship between the two cutoffs and equation (13), I can determine the equilibrium price of non-manufacturing:  $p_s = 1 + \frac{p_o O}{e}$ . A higher windfall translates into a higher service sector price which results in

<sup>16</sup> Whilst I focus on heterogenous workers, this setup can easily be related to one with heterogenous firms without changing the results.

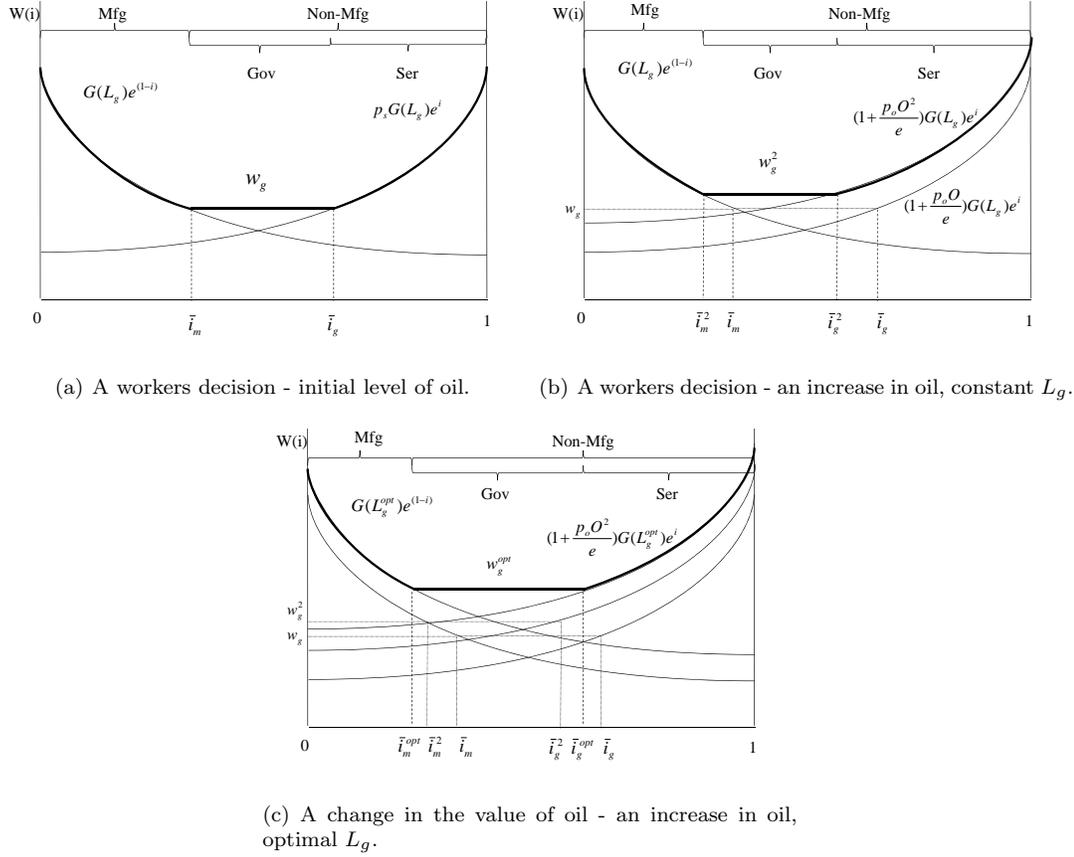


Figure 1: The mechanics of the model in a simple example.

an increase in service sector wage offers. In order for employment in government to remain unchanged despite the higher price, wages in the government sector must also rise. This results in a shift of workers from manufacturing to government and from government towards services resulting in a leftward shift of both cutoffs to  $\bar{i}_m^2(p_s^2; L_g^s)$  and  $\bar{i}_g^2(p_s^2; L_g^s)$ . As both cutoffs shift left, manufacturing productivity ( $Y_m/\bar{i}_m = (e - e^{1-\bar{i}})/\bar{i}_m$ ) rises: the workers who remain in the manufacturing sector are most skilled in manufacturing sector work. At the same time service sector productivity ( $Y_s/(1 - \bar{i}_s) = (e - e^{\bar{i}})/(1 - \bar{i}_s)$ ) falls: new entrants in non-manufacturing pull down productivity since they are, on average, less skilled than those already employed in non-manufacturing. Finally, I can also show that non-manufacturing sector productivity,  $Y_s/(1 - \bar{i}_s + L_g) = (e - e^{\bar{i}})/(1 - \bar{i}_s + L_g)$  also falls as long as the government sector is not “too large”.<sup>17</sup>

<sup>17</sup> In particular  $L_g < -2 - \log(e + p_o O) + 2\Omega(e + p_o O)$ , where  $\Omega(\cdot)$  is the product logarithm function.

**Government** So far, I have taken the size of government employment as fixed. Suppose however that the fiscal authority takes the demand functions of agents (derived above in equations (12)) as given and solves the Ramsey-type problem for the optimal size of government employment,  $L_g^{opt}$ , in equation (14). Taking the first order condition from this maximization problem, and applying the implicit function theorem to the resulting first order condition, it can be shown that the optimal size of government increases with the size of the oil endowment:  $\frac{\partial L_g^{opt}}{\partial p\omega O} > 0$ . Intuitively, higher oil endowment means a greater demand for both traded and non-traded goods. Demand for traded, manufacturing goods can be satiated by imports from abroad. Demand for non-traded goods however - which include government services - is satiated with locally produced goods, and hence results in a shift of labor towards the non-traded sectors of services and government. This is shown in Figure 1(c). The impact on manufacturing productivity is unambiguous - manufacturing productivity will increase both due to the smaller size of the manufacturing sector (and thus its more specialized nature) as well as due to the larger government sector which in turn increases each workers productivity. The impact on non-manufacturing productivity is mixed and will depend on specific parameters - but can potentially be negative.

## 5 Solving the Model

**Distribution Function** To calibrate and solve the model, I must pick a particular parametric form for the distribution of skills,  $N(z_s, z_m)$ , since the Roy model cannot be identified from cross-sectional wage data alone.<sup>18</sup> In what follows, I assume that skills are drawn independently from a normalized Type II extreme value (or Frechet) distribution with CDF:

$$N(z_s) = e^{-z_s^{-\theta}} \text{ and } N(z_m) = e^{-z_m^{-\theta}}, \quad (15)$$

where,  $\theta > 1$ . The log of a random talent draw,  $\log Z_i$ , has standard deviation  $\pi/(\theta\sqrt{6})$ , where  $\pi$  is the constant. The parameter  $\theta$  thus governs the amount of variation in skills and hence the observed productivity dispersion: lower values of  $\theta$  imply more heterogeneity in ability and higher productivity dispersion. Notice that I assume that  $\theta$  is common to both manufacturing and service sectors and that talent draws are independent of each other. Whilst both these assumptions may seem restrictive, they allow me to derive simple, analytic solutions which provide insights into the workings of the model. In Kuralbayeva and Stefanski (2013) we had extend the (no-government version) of the model to allow correlated talent draws and different dispersions across sectors and we had shown that, quantitatively, these channels played only a limited role.

---

<sup>18</sup> This is because we observe only the outcomes of workers choices (in the form of a worker's observed wages) and not the talent draws (and hence the sectoral wage offers) that underpin these outcomes.

I focus on the Frechet distribution for several reasons. First and foremost, this distribution is one of three extreme value distributions. According to the Fisher - Tippet - Gnedenko theorem from extreme value theory, there are only three types of distributions that are needed to model the maximum or minimum of the collection of random observations from the same distribution. More specifically, the maximum of a sample of i.i.d. random variables converges in distribution to either the Gumbel, the Frechet, or the Weibull distribution.<sup>19</sup> In my case, choosing an extreme value distribution can be thought of as capturing the distribution of agents' "best" talents in each particular sector. Second, of these three distributions I choose the Frechet in keeping with the literature. Obviously, Kuralbayeva and Stefanski (2013) chose this distribution. Furthermore, Eaton and Kortum (2001) have used this distribution to parameterize a Ricardian model of international trade and Lagakos and Waugh (2012) have used it to model talent distribution across sectors. Finally, the Frechet distribution also provides very tractable analytic solutions which allow for easy interpretation of results and does a very good job of fitting the data.

**Employment** Since  $z_s$  and  $z_m$  are independently drawn from Frechet distribution, the joint density function can be expressed as  $g(z_s, z_m) = g(z_s)g(z_m)$ . Using this, I can relate sectoral labor supply allocation to the parameter which controls the dispersion of skills across sectors.

First, start with government employment. In order to induce  $L_g$  workers to work in the government sector, the government will have to offer a wage,  $w_g$ , such that enough workers are drawn to that sector by earning more than they could in either manufacturing or non-manufacturing. Consequently, the chosen wage will be defined by the following:

$$L_g = P(w_g > w_m, w_g > w_s) = P(w_g > G_m z_m, w_g > p_s G_s z_s) = \int_0^{\frac{w_g}{G_s p_s}} \int_0^{\frac{w_g}{G_m}} g(z_s)g(z_m) dz_m dz_s. \quad (16)$$

Taking the level of government employment (and hence government wage) as given, the expected employment in services and manufacturing is:

$$L_s = P(w_s > w_m, w_s > w_g) = P(p_s G_s z_s > G_m z_m, p_s G_s z_s > w_g) = \int_{\frac{w_g}{G_s p_s}}^{\infty} \int_0^{p_s G_s z_s / G_m} g(z_s)g(z_m) dz_m dz_s \quad (17)$$

---

<sup>19</sup> Broadly speaking, if one generates  $N$  data sets from the same distribution, and then creates a new data set that includes only the maximum values of these  $N$  data sets, the resulting data set can only be described by one of the above distributions. For more details see Haan and Ferreira (2006).

$$L_m = P(w_m > w_s, w_m > w_g) = P(G_m z_m > p_s G_s z_s, G_m z_m > w_g) = \int_{w_g/G_m}^{\infty} \int_0^{\frac{G_m z_m}{p_s G_s}} g(z_s)g(z_m)dz_s dz_m \quad (18)$$

Given the Frechet distribution of talent draws, the above equations can be simplified into the following expressions, which depend only on the given level of government employment and the price of non-manufacturing goods:

$$L_s = \frac{G_s^\theta p_s^\theta}{G_m^\theta + G_s^\theta p_s^\theta} (1 - L_g), \quad L_m = \frac{G_m^\theta}{G_m^\theta + G_s^\theta p_s^\theta} (1 - L_g), \quad L_g = e^{-\frac{G_m^\theta + G_s^\theta p_s^\theta}{w_g^\theta}}. \quad (19)$$

Solving this for  $w_g$  I obtain:

$$w_g = \left( -\frac{G_m^\theta + G_s^\theta p_s^\theta}{\log(L_g)} \right)^{\frac{1}{\theta}}. \quad (20)$$

**Output** Normalizing  $A = 1$ , the output of each sector can be expressed as:

$$Y_s = G_s \int_{\frac{w_g}{G_s p_s}}^{\infty} \int_0^{\frac{p_s G_s z_s / G_m}{w_g / G_m}} z_s g(z_s, z_m) dz_m dz_s, \quad Y_m = G_m \int_{w_g / G_m}^{\infty} \int_0^{\frac{G_m z_m}{p_s G_s}} z_m g(z_s, z_m) dz_s dz_m \quad (21)$$

Since  $z_s$  and  $z_m$  are independently drawn from a Frechet distribution, this simplifies to:

$$Y_s = G_s \left( \frac{G_s^\theta p_s^\theta}{G_s^\theta p_s^\theta + G_m^\theta} \right)^{1 - \frac{1}{\theta}} \Lambda(\theta, L_g), \quad Y_m = G_m \left( \frac{G_m^\theta}{G_s^\theta p_s^\theta + G_m^\theta} \right)^{1 - \frac{1}{\theta}} \Lambda(\theta, L_g). \quad (22)$$

where  $\Lambda(\theta, L_g) \equiv \Gamma(1 - \frac{1}{\theta}) - \Gamma(1 - \frac{1}{\theta}, -\log(L_g))$ , whilst  $\Gamma(\cdot)$  and  $\Gamma(\cdot, \cdot)$  denote the complete and incomplete gamma functions.

For a given level of government employment,  $L_g$ , using the above equations for sectoral output and 13, it is easy to show that  $\frac{\partial p_s}{\partial p_o O} > 0$ . It then follows that oil endowments result in a reallocation of labor:  $\frac{\partial L_s}{\partial p_o O} > 0$  and  $\frac{\partial L_m}{\partial p_o O} < 0$ . This shift in labor generates specialization (in manufacturing) and de-specialization (in services):  $\frac{\partial Y_s / L_s}{\partial p_o O} < 0$  and  $\frac{\partial Y_m / L_m}{\partial p_o O} > 0$ . If - instead - I consider productivity in the non-manufacturing sector, I can also show that  $\frac{\partial Y_s / (L_s + L_g)}{\partial p_o O} < 0$  as long as government employment is not ‘too-large’ i.e. if and only if  $L_g < \frac{1}{\theta - 1} \frac{L_s}{L_s + L_m}$ . Later, in the calibration, it is easy to verify that this condition is satisfied for every country-date in our data-set.

## 6 Calibrating the Model

**Estimating Skill Dispersion** The parameter  $\theta$  governs the dispersion of (unobserved) underlying skills. To match this parameter to observed variables, I make use of the properties

of the Frechet distribution. The distribution of wage offers in each (non-government) sector is given by:

$$N_s^w(w_s) = Pr\{W_s \leq w_s\} = Pr\{p_s AG_s Z_s \leq w_s\} = Pr\{Z_s \leq \frac{w_s}{p_s AG_s}\} = e^{-(p_s AG_s)^\theta w_s^{-\theta}} \quad (23)$$

$$N_m^w(w_m) = Pr\{W_m \leq w_m\} = Pr\{AG_m Z_m \leq w_m\} = Pr\{Z_m \leq \frac{w_m}{AG_m}\} = e^{-(AG_m)^\theta w_m^{-\theta}}. \quad (24)$$

These are both Frechet density functions with the same dispersion parameter,  $\theta$ , as the talent distributions.<sup>20</sup> Thus, the wage offers in the non-governmental sector is the maximum an agent could earn in either sector,  $w_{ng} = \max\{w_s, w_m\}$ . The distribution of this wage,  $N^{ng}(w)$ , is then the maximum order statistic of wage offers in non-governmental sectors and is given by:

$$N^{ng}(w) = N_s^w(w)N_m^w(w) = e^{-A^\theta (G_m^\theta + G_s^\theta p_s^\theta) w^{-\theta}}. \quad (25)$$

The above distribution is also Frechet with the same dispersion parameter (but with a different mean) as the skill distribution. This is a consequence of the assumption that the original talents were drawn from an extreme value distribution. Finally, agents with a non-governmental wage offer drawn from this distribution will choose to work in government if and only if the wage offered by government,  $w^g$ , is higher than their non-governmental wage offer. Consequently, the distribution of observed wages will be given by:

$$N(w) = N^{ng}(w)\mathbb{1}_{w \geq w_g}(w). \quad (26)$$

In order to match the parameter  $\theta$ , I use a method of moments. In particular I calculate the standard deviation of a sample of log wages in a ‘resource poor’ country and match it to the implied standard deviation of log wages in the model. As in Kuralbayeva and Stefanski (2013) I obtain cross-sectional wage data from the 2009 US Current Population Survey (CPS) and find that the standard deviation of log wages in this sample is 0.58.<sup>21</sup> Then, I calculate the corresponding theoretical standard deviation of the log-wage and choose  $\theta = 2.43$  so that the two match.<sup>22</sup>

<sup>20</sup> Notice that these are not distributions of observed wages in a given sector, but the distribution of (unobserved) wages that agents could earn if they chose to work in a particular sector.

<sup>21</sup> Following Kuralbayeva and Stefanski (2013), Lagakos and Waugh (2012) and Heathcote et al. (2009) I include individuals aged 25 to 60 who have non-missing data on income and hours worked. Wages are before tax, and are taken to be the sum of wage, business and farm income. The sample is further restricted to include workers who average more than 35 hours a week of work and earn at least the Federal minimum wage.

<sup>22</sup> To do this, notice that for any integrable function,  $f$ , I can write:  $\mathbb{E}(f(w)) = f(w_g) \int_0^{w_g} dG^{ng}(w) + \int_{w_g}^{\infty} f(w) dG^{ng}(w)$ . Noting that the standard deviation of log-wages in the model,  $\sigma$ , is given by  $\sigma = \sqrt{Var(\log(W))} = \sqrt{\mathbb{E}(W^2) - (\mathbb{E}(W))^2}$ , I use the above formula and the CDF of  $N^{ng}(w)$  to calculate  $\theta$ .

**Government parameters** To calibrate the government parameters, I first impose the restriction that  $\psi \equiv \psi^s = \psi^m$  so that the impact of government on productivity is the same in both manufacturing and service sectors. The reason for this assumption is twofold: first, it simplifies the analysis and second there is no a priori reason to believe that the impact of government spending should affect productivity more in one sector than in another. Having made this assumption, I choose  $\psi$  so that the predicted *optimal* government employment in resource poor countries in the model exactly matches government employment in the lowest decile of resource exporting countries in the data of approximately 17%. Thus, notice I am purposefully assuming resource poor countries have efficient levels of government in order to obtain the lower bound measures of government inefficiency in resource-rich countries. In other words, if resource poor countries have inefficient government, then the scale of misallocation in resource-rich countries will be even larger than the model suggests. Thus, my measure will provide a lower bound or best-case-scenario for the extent of misallocation in resource-rich countries.

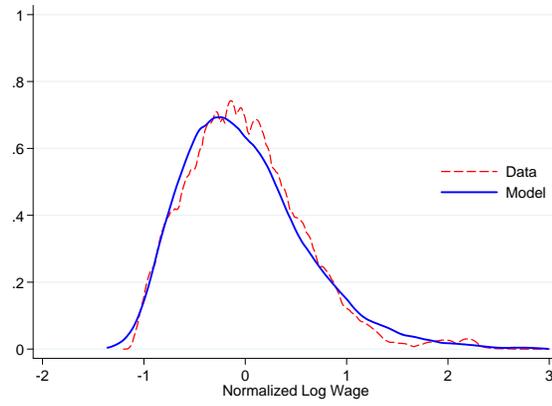
**Preference parameters** Finally, I follow Kuralbayeva and Stefanski (2013) in estimating preference parameters  $\sigma$  and  $\nu$ . From the household's problem I can derive an equation relating relative consumer expenditure on the relative price:  $\frac{c_m}{c_s} = (\nu p_s)^\sigma$ . Taking logs of this equation, I estimate elasticity of substitution between manufacturing and non-manufacturing goods using ICP data and find that  $\sigma = 0.94$ . Finally, I choose the preference parameter to be  $\nu = 0.29$ , to match the employment share in the manufacturing sector in resource-poor countries in the model to the employment share in manufacturing in the lowest decile of exporters in our sample (approximately 19%).

## 7 Results

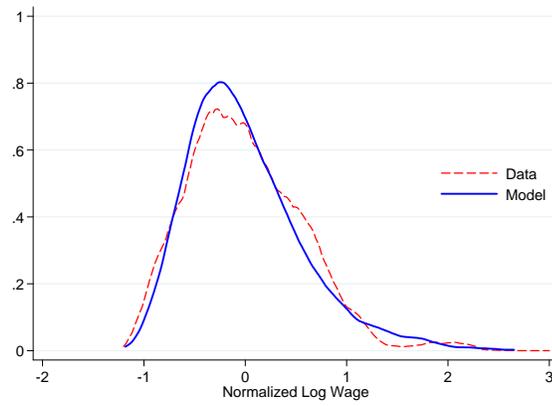
**US Wage Distributions** In the calibration, an important parameter choice is the dispersion parameter of the Frechet distribution,  $\theta$ , chosen to match the standard deviation of observed log wages. Figure 2, shows the theoretical and empirical density of observed wages at the aggregate and sectoral level in the US data and the model. The Frechet distribution does well at matching the general dispersion and the fat tails of wages at the aggregate level.<sup>23</sup> This finding mirrors that of Kuralbayeva and Stefanski (2013), Lagakos and Waugh (2012) and Heckman and Sedlacek (1985). The model with Frechet talent draws also does well in matching the dispersion and the tails of the data that we have not directly calibrated: the observed sectoral wages in the US.

---

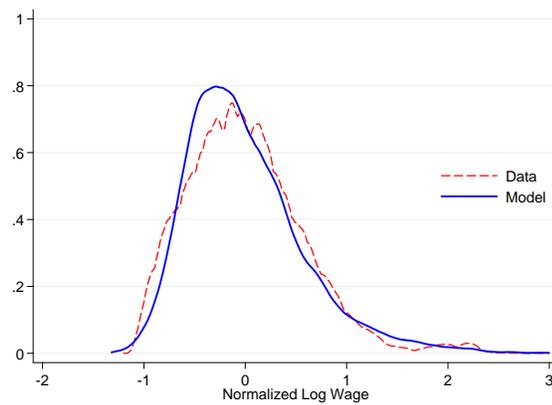
<sup>23</sup> This is unlike, a log-normal distribution, for instance, which fails to match the fat tails - see Kuralbayeva and Stefanski (2013).



(a) Distribution of Wages, Frechet



(b) Distribution of Wages in Mfg, Frechet



(c) Distribution of Wages in Non-Mfg., Frechet

Figure 2: Distribution of wages in baseline model and data, by sector.

	Data	Model			Model/Data		
		(1)	(2)	(3)	(1)	(2)	(3)
		Opt. Gov.	Obs. Gov.	No Gov.	Opt. Gov.	Obs. Gov.	No Gov.
M. Emp., $L_m$	-0.014	-0.008	-0.016	-0.010	0.55	1.14	0.67
M. Prod, $D_m$	0.068	0.022	0.025	0.030	0.32	0.37	0.44
NM. Prod, $D_s$	-0.012	-0.004	-0.008	-0.007	0.30	0.60	0.59
NM. Price, $p_s$	0.048	0.027	0.027	0.037	0.56	0.56	0.76
G. Emp., $L_g$	0.017	0.002	0.017		0.13	1.00	-

Table 4: Changes in sectoral employment and sectoral productivity associated with resource wealth in the Data and Model.

**Productivity and Employment** Next, to test the success of the model on cross country data, I consider three different versions of the model: 1) a model where government employment is chosen optimally; 2) a model where government employment is taken directly from the data; and 3) a model without government.

Table 4 compares the empirical windfall elasticities from the data (shown in Tables 2 and 3) with the corresponding windfall elasticities implied by the different versions of the model for sectoral employment, productivity and prices.<sup>24</sup> First, in column (1) of Table 4, I examine how the elasticities in the data compare to the model where government employment shares are chosen optimally. A doubling of natural resource windfalls in the optimal model results in a 0.8 percentage point decline in manufacturing employment, a 2.2% increase in manufacturing productivity, a 0.4% decline in non-manufacturing productivity and a 2.7% increase in the price of non-manufacturing goods. With respect to these measures the model does relatively well, explaining between 30% and 56% of the observed changes. Where the optimal model does very poorly is in explaining the elasticity of government employment. Here a doubling of windfalls in the model is associated with only 0.2 percentage point increase in government employment - whereas in the data a doubling of windfalls is associated with a 1.7 percentage point increase. Thus, the model explains only 13% of the observed elasticity. This suggests that resource-rich countries have a much higher government employment share than they the model predicts that they “should”.

Next, in column (2) of Table 4, instead of considering the optimal government employment, I instead feed into the model the observed government employment shares. Now, the performance of the model improves dramatically across all measures. The model captures all - and even slightly over-predicts - the elasticity of manufacturing employment. It also explains 37% of

<sup>24</sup> In the data, we measure resource wealth as the ratio of current price exports of natural resources to current price GDP measured in international dollars. International dollar are constructed to have the same purchasing power over GDP as the U.S. dollar has in the United States. Since the US is a resource-poor country (according to this measure), we can view GDP in international dollars as the GDP of a country measured using a resource-poor country’s prices. As such, in the model, we construct our resource wealth measure as the value of exports of natural resources divided by GDP, measured with the prices of a resource-poor country (i.e. one that has  $p_O = 0$ ).

the elasticity in manufacturing productivity, 60% of the productivity in non-manufacturing productivity and 56% of the elasticity of non-manufacturing prices and - by construction - it accounts for all of the government employment elasticity.

Finally, in column (3) of Table 4, I consider a no-government version of the model.<sup>25</sup> Notice, that the non-government version of the model is approximately as successful in explaining most of the observed facts as the observed-government version of the model in column (2)- with the important exception that it does not take into account the employment share of the government.

The message from this exercise is that the specialization mechanism introduced in Kuralbayeva and Stefanski (2013) is strong enough to explain a big part of the large differences in sectoral employment shares and asymmetric productivity differences between resource-rich and resource-poor countries. Furthermore, the differences in the size of government employment between resource-rich and resource poor countries act to magnify the differences in sectoral productivity and employment produced by the specialization effect. Most importantly, the observed government employment shares in resource-rich countries, tend to be significantly “too-large”. I explore the impact of this latter effect on welfare and productivity in the following section.

## 8 The Resource Curse

The above results suggest one particular mechanism that fits the description of the so-called resource-curse - a well-known, stylized fact relating negative economic outcomes to resource windfalls. In the context of this paper, the mechanism for a resource curse is clear. If there is a misallocation of public sector employment in resource rich countries - thus government employment is either too large or too small relative to the optimum - this misallocation of workers across sectors and a lower productivity and welfare in the model. Table 5 shows the regressions of the ratio of observed-to-optimal aggregate productivity and welfare respectively emerging from the model, versus the size of natural resource windfalls (and the log of natural resource windfalls). From columns (1) and (2) observe that - in the data - a doubling of the natural resource windfall is associated with productivity that is 0.7% lower and a welfare that is 0.5% lower than it otherwise could be, if government employment were not mis-allocated. In other words, from columns (3) and (4), a one percentage point increase in resource export shares in GDP is associated with a productivity that is 0.19% lower than it otherwise could be and a welfare that is 0.12% lower than it otherwise could be.

As I showed before, the misallocation occurs due to a government sector that tends to be too large in resource-rich countries. Importantly, I make absolutely no claims as to *why* the size of

---

<sup>25</sup> Thus, here the model is re-calibrated in that  $\psi$  is set to zero, and all other parameters are chosen to match the relevant moments described in the paper. In particular, I choose  $\nu = 0.225$ ,  $\theta = 2.22$  and  $\sigma = 0.94$ .

government employment tends to be what it is and in particular why government employment tends to be higher in resource-rich countries. Government employment in resource-rich countries can be non-optimal for a whole host of reasons (some associated with natural resources, and others not), but this paper takes no stand on the issue, and simply takes the observed size of government in resource-rich countries as given. As such, the observation that resource-rich countries have larger than optimal government is a characteristic of the given sample of data, and will not necessarily hold in every single resource-rich country. The findings here thus reflect the fact that in this particular sample of data, resource-rich countries tended to have public employment that was “too-large”. It is - of course - entirely possible to find examples of resource-rich countries in the sample which the model predicts had “too-small” or “just-right” government. As two such specific examples, consider the cases of Chile and Canada in 2007. Chile’s windfall measure was approximately 20% of GDP. This was associated with a productivity that was approximately 3% lower and a welfare that was approximately 4% lower than it otherwise could have been. This lower productivity and welfare, was a consequence of a government sector employment share that was - according to the above model - approximately 8.5 percentage points too *small* relative to predicted optimum. In the case of Canada, its windfall measure was approximately 11% of GDP in 2007. This was associated with a productivity that was only 0.01% lower and a welfare that was approximately 0.02% lower than it otherwise could have been. This was a consequence of the fact that Canada almost had ‘the right’ levels of government employment given its resource windfall.

The above view fits in well with institutional view of the resource curse. In particular by emphasizing the role of government misallocation, my theory lends support to arguments by Robinson et al. (2006), van der Ploeg (2010) and others that explanations of the resource curse should be sought outside economic structure perhaps - as they suggest - in areas such as political economy, weak institutions or property rights.

## 9 Conclusion

Kuralbayeva and Stefanski (2013), show that - in the data - resource-rich regions have small and productive manufacturing sectors and large and unproductive non-manufacturing sectors and propose a mechanism that explains these productivity differences through a process of self-selection. Windfall revenues induce labor to move from the (traded) manufacturing sector to the (non-traded) non-manufacturing sector. A self-selection of workers takes place. Only those most skilled in manufacturing sector work remain in manufacturing. Workers that move to the non-manufacturing sector however, are less skilled at non-manufacturing sector work than those who were already employed there. Resource-induced structural transformation thus results in

	(1) Rel. Prod. ( $D^{obs}/D^{opt}$ )	(2) Rel. Welf. ( $U^{obs}/U^{opt}$ )	(3) Rel. Prod. ( $D^{obs}/D^{opt}$ )	(4) Rel. Welf. ( $U^{obs}/U^{opt}$ )
log(NRE)	-0.007*** (0.001)	-0.005*** (0.001)		
NRE			-0.194*** (0.025)	-0.122*** (0.024)
Time FE	yes	yes	yes	yes
Obs.	340	340	340	340
$R^2$	0.114	0.085	0.182	0.106

Standard errors in parentheses  
\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Table 5: Regressions of the ratios of productivity and welfare in the observed and optimal models with respect to resource wealth.

higher productivity in manufacturing and lower productivity in non-manufacturing.

In this paper, I show that in addition to the above facts, in the data, resource-rich countries also tend to employ a larger proportion of workers in the government sector than resource-poor countries. I then adapt the model of specialization of Kuralbayeva and Stefanski (2013) to include a productive government sector and proceed to examine optimal government employment in resource-rich countries. In particular, I show that the model can generate higher employment in the government sector when windfalls are higher. In a nutshell, government services are non-traded. Higher windfalls will increase demand for all goods and services - including government services - but since these cannot be imported, workers will shift to the government sector to satiate demand. Furthermore, even with a government sector, the specialization mechanism introduced in Kuralbayeva and Stefanski (2013) is strong enough to explain a large part of the asymmetric differences in sectoral employment shares and productivity between resource-rich and resource-poor countries. In addition, the differences in the size of government between resource-rich and resource poor countries act to magnify the differences in sectoral productivity and employment shares produced by this specialization mechanism. Finally, the observed government employment shares in resource-rich countries, tend to be “too-large” relative to optimum. In the calibrated best-case-scenario model, government employment is nearly 10 times smaller than in the data. This implicit misallocation of resources has a large, negative impact on welfare and aggregate productivity. Using the calibrated model I find that a ten percentage point increase in resource windfalls is associated with a 1.94% percent lower aggregate productivity and a 1.22% lower welfare arising from government misallocation in resource-rich countries.

As such, the above theory and empirical evidence suggest that institutions may play a key role in driving the resource curse. In particular, this paper lends support to arguments by Robinson et al. (2006), van der Ploeg (2010) and others that explanations of the resource curse

should be sought outside economic structure perhaps - as they suggest - in areas such as political economy, weak institutions or property rights which induce governments to be particularly large in resource-rich countries.

## 10 Data Appendix

**Resource Wealth** We follow Sachs and Warner (2001) in defining natural resource “wealth” as the ratio of exports of natural resources (fuels, ores and metals) to GDP using WDI data. Unlike Sachs and Warner (2001), we use PPP GDP (in current prices) in the denominator of our measure. We do this since higher endowments of resources can potentially impact prices of non-resource goods (and hence measured GDP) influencing both the numerator and the denominator of our measure. Using PPP GDP, keeps prices fixed across countries and hence the measure only captures changing resource wealth. We have experimented with both measures of resource wealth, as well as other measures such as the ratio of *net* exports of natural resources to gross domestic product (both observed price and PPP). Our results, however, are unaffected.

**Labor Shares** To calculate the last two measures of productivity, we need to find expressions for labor shares,  $1 - \alpha_s$ , for each sector  $s$ . Although these shares can potentially vary across countries, due to a lack of comprehensive cross-country sectoral data, we make use of OECD data to calculate the average annual share of employee compensation for each sector in OECD countries for the longest period of time that data is available. We calculate the labor share as the ratio of total compensation of employees (wages and salaries before taxes, as well as employer’s social contributions) over sectoral value-added.<sup>26</sup> Table 6 presents the results. We find labor share in manufacturing is 0.57 whilst in non-manufacturing it is 0.53. Notice that these are lower bound estimates since national accounts data does not include incomes generated from self-employment under total compensation. One commonly used technique to correct for this is to re-scale the shares by the ratio of total employment to total employees. In effect, the self-employed are then assumed to be paid the same average rate of compensation as employees and the same marginal rate of productivity is assumed for dependent and independent workers. Doing this with the above data results in higher average labor shares of 0.64 in non-manufacturing and 0.62 in manufacturing. Quantitatively and qualitatively however, this adjustment leaves our results almost unchanged. Since we are uncertain of exactly how compensation of self-employed workers varies across countries and sectors, we choose to retain our first estimates of labor shares.

**Sectoral Employment** We obtain sectoral employment data for 1980-2006 from the ILO KILM online database. To obtain the largest set of employment data, we combine ISIC revision 2 and ISIC revision 3 employment data.<sup>27</sup>

<sup>26</sup> Tables 7 and 8 in the the OECD Annual National Accounts, Volume 2, 1970-2008 (2009 prov)- Detailed aggregates, in millions of national currency

<sup>27</sup> Since these sector classifications are at the one digit level, there are no issues with the correspondence between ISIC Rev.3 and Rev.2. We use the rule that if data is available in both revisions, we use revision 3 data. Finally, we do not consider employment data based entirely on urban surveys - as these significantly underestimate employment in agriculture and overestimate employment in other sectors.

Country	Period	Labor Shares							
		Agr	Cons.	Ser	Mfg	M&U	ACS	ACSM	Total
Australia	1989-2008	0.26	0.49	0.58	0.55	0.29	0.56	0.56	0.54
Austria	1976-2008	0.11	0.62	0.59	0.67	0.43	0.57	0.60	0.59
Belgium	1995-2008	0.17	0.57	0.57	0.65	0.38	0.56	0.58	0.57
Canada	1981-2006	0.31	0.70	0.60	0.61	0.24	0.59	0.60	0.57
Chile	2003-2008	0.40	0.65	0.51	0.30	0.10	0.52	0.48	0.40
Czech.	1995-2008	0.46	0.50	0.47	0.52	0.33	0.47	0.48	0.48
Denmark	1970-2008	0.26	0.79	0.62	0.72	0.21	0.62	0.63	0.62
Finland	1975-2008	0.19	0.69	0.63	0.60	0.34	0.60	0.60	0.59
France	1999-2008	0.22	0.57	0.58	0.66	0.41	0.57	0.58	0.58
Germany	1991-2008	0.43	0.66	0.53	0.72	0.50	0.54	0.58	0.58
Greece	2000-2008	0.12	0.35	0.40	0.49	0.33	0.38	0.39	0.39
Hungary	1995-2008	0.31	0.51	0.54	0.54	0.50	0.52	0.53	0.53
Iceland	1997-2005	0.61	0.57	0.67	0.70	0.29	0.65	0.66	0.64
Ireland	1995-2008	0.18	0.68	0.53	0.30	0.47	0.53	0.46	0.46
Italy	1970-2008	0.30	0.44	0.47	0.59	0.40	0.46	0.49	0.49
Japan	1996-2007	0.24	0.72	0.50	0.53	0.28	0.51	0.52	0.51
Korea	1970-2008	0.11	0.64	0.49	0.50	0.31	0.44	0.46	0.46
Lux.	1995-2008	0.26	0.67	0.50	0.63	0.35	0.51	0.52	0.52
Mexico	2003-2007	0.16	0.39	0.33	0.30	0.11	0.33	0.32	0.30
Nlands	1970-2008	0.21	0.72	0.62	0.64	0.17	0.60	0.61	0.59
New Zeal.	1986-2006	0.22	0.51	0.47	0.53	0.19	0.45	0.46	0.45
Norway	1970-2009	0.22	0.74	0.62	0.73	0.15	0.61	0.63	0.54
Poland	1995-2008	0.20	0.41	0.42	0.56	0.55	0.40	0.43	0.44
Portugal	1995-2006	0.18	0.63	0.59	0.62	0.30	0.57	0.58	0.57
Slovakia	1995-2008	0.43	0.39	0.44	0.47	0.33	0.44	0.44	0.44
Spain	1995-2008	0.19	0.62	0.54	0.61	0.28	0.53	0.54	0.54
Sweden	1993-2008	0.31	0.76	0.64	0.63	0.25	0.64	0.63	0.62
US	1987-2007	0.29	0.67	0.58	0.64	0.27	0.58	0.59	0.58
Average		0.26	0.59	0.54	0.57	0.31	0.53	0.53	0.52

Table 6: Shares of labor compensation relative to sectoral value added in agriculture, construction, services, manufacturing, mining and utilities as well as (non-resources) non-manufacturing (which consists of agriculture, construction and services), the non-resource sector (agriculture, construction, services and manufacturing) as well as total value added (agriculture, construction, services, manufacturing, mining and utilities) in OECD countries for the periods indicated. (Source: OECD, 2007)

**Prices** Since we want to compare sectoral productivity across countries, it is crucial to control for any price differences that may exist between sectors across countries. We do this by constructing country- and sector-specific price levels for each sector. In particular, we use the World Bank's 2005 International Comparison Program (ICP) database which provides cross-country data on the value of final household and government expenditures by sector for the year 2005. Expenditure data is given in current US dollars (at market exchange rates), as well as in PPP terms which allows us to extract country specific sectoral price levels (relative to the corresponding price level in the US). Denoting current price and PPP expenditures on sector  $s$  goods in country  $i$  by  $E_s^i$  and  $E_s^{PPP}$  respectively, the price level of sector  $s$  in country  $i$  (relative

<b>Agriculture</b>	<b>Manufacturing</b>	<b>Construction</b>	<b>Services</b>	<b>Min. &amp; Util.</b>	<b>Not Otherwise Class.</b>
1101112 Other cereals & flour	1103111 Clothing mat. & access.	150210 Res. buildings	1103141 Clean. & repair of clothing ser.	110440 Water supply & misc. dwelling ser.	111300 Net purchases abroad
1101113 Bread	1103121 Garments	150220 Non-res. buildings	1103221 Repair & hire of footwear	110451 Electricity	130221 Comp. of empl.
1101114 Other bakery products	1103211 Footwear	150230 Civil eng. works	110410 Rentals for housing	110452 Gas	130222 Intern. cons.
1101115 Pasta products	110511 Furniture & furnishings		110430 Maint./repair of the dwelling	110453 Other fuels	130223 Gross oper. surplus
1101121 Beef & veal	110512 Carpets & other floor coverings		110442 Misc. ser. relating to the dwelling	110722 Fuels/lubes for pers. transp. equip.	130224 Net taxes on production
1101122 Pork	110520 Household textiles		110513 Repair of furniture	110733 Pass. trans. by air	130225 Receipts from sales
1101123 Lamb, mutton & goat	110531 Major HH apps.		110533 Repair of HH apps.	110734 Pass. trans. by sea & inland waterway	130421 Comp. of empl.
1101124 Poultry	110532 Small electric HH apps.		1105621 Domestic ser.	110735 Comb. passenger trans.	130422 Intern. cons.
1101125 Other meats & preparations	110540 Glassware, tableware & HH utensils		1105622 Household ser.	110736 Other purch. transp. ser.	130423 Gross operating surplus
1101131 Fish & seafood	110551 Major tools & equip.		110621 Medical ser.	110810 Postal ser.	130424 Net taxes on prod.
1101132 Pres. fish & seafood	110552 Small tools & miscellaneous access.		110622 Dental ser.	110830 Tel. & telefax ser.	130425 Receipts from sales
1101141 Fresh milk	110561 Non-durable HH goods		110623 Paramedical ser.		140111 Comp. of empl.
1101142 Pres. milk & milk products	110611 Pharmaceutical products		110630 Hospital ser.		140112 Intern. cons.
1101143 Cheese	110612 Other medical products		110723 Maint. & repair of pers. transp. equip.		140113 Gross op. surplus
1101144 Eggs & egg-based products	110613 Therap. apps. & equip.		110724 Other ser. in respect of pers. transp. equip.		140114 Net taxes on production
1101151 Butter & margarine	110711 Motor cars		110731 Pass. trans. by rail		140115 Receipts from sales
1101153 Other edible oils & fats	110712 Motor cycles		110732 Pass. trans. by road		160000 Change in inv. & valuables
1101161 Fresh or chilled fruit	110713 Bicycles		110733 Pass. trans. by air		180000 Balance of ex. & im.
1101162 Frozen, pres. or processed fruits	110820 Tel. & telefax equip.		110734 Pass. trans. by sea & inland waterway		
1101171 Fresh or chilled vegetables	110911 AV, phot.& info. proc. equip.		110735 Comb. passenger trans.		
1101172 Fresh or chilled potatoes	110914 Recording media		110736 Other purch. transp. ser.		
1101173 Frozen or pres. vegetables	110921 Major durables for outdoor & indoor recreation		110810 Postal ser.		
1101181 Sugar	110931 Other recreational items & equip.		110915 Repair of AV, phot.& info. proc. equip.		
1101182 Jams, marmalades & honey	111212 Apps., articles & products for pers. care		110933 Gardens & pets		
1101183 Confectionery	111231 Jewellery, clocks & watches		110935 Vet. ser. for pets		
1102111 Spirits	111232 Other pers. effects		110941 Rec. & sporting Cultural ser.		
1102121 Wine	150110 Metal products & equip.		110942 Cultural ser.		
1102131 Beer	150120 Transport equip.		110943 Games of chance		
110119 Food prod. n.e.c.	150300 Other products.		110950 News., books & stationery		
110121 Coffee, tea & cocoa			110960 Package holidays		
110122 Mineral waters, soft drinks, fruit & veg. juices			111000 Education		
110220 Tobacco.			111100 Catering ser.		
			111120 Accom. ser.		
			111211 Hairdressing salons & pers. grooming est.		
			111220 Prostitution		
			111240 Social protection		
			111250 Insurance		
			111261 FISIM		
			111262 Other financial ser.		
			111270 Other ser. n.e.c.		

Table 7: ICP Disaggregated Categories

to that of the US) is given by:

$$P_s^i/P_s^{PPP} = E_s^i/E_s^{PPP}. \quad (27)$$

The publicly available ICP expenditure data disaggregates expenditure into 19 sectors. These however, do not map very well into ISIC sectors. On request however, it is possible to obtain proprietary ICP data that is further disaggregated into approximately 129 sectors. We make use of this disaggregated data to construct expenditure data at market exchange rates and at PPP for five sectors: agriculture, manufacturing, mining & utilities, construction and services. Our mapping of ICP to ISIC data is shown in Table 7. Notice that the last columns refers to categories that are not classified and hence excluded from our price indices. We then use equation 27 to calculate sector specific price levels in each country (relative to the US) for each of the five sectors. Finally, we can combine the above 2005 price level data with sectoral price indices for the 1980-2008 period from the UN to obtain a panel of price *levels*. In particular, we obtain one digit ISIC v.3 sectoral value-added data in constant 1990 USD prices ( $VA_{s,t}^{1990}$ ) and current prices ( $VA_{s,t}$ ) which can be used to calculate an index of sectoral prices relative to a base year:  $P_{s,t}/P_s^{1990} = VA_{s,t}/VA_{s,t}^{1990}$ . We then rebase this index so that it is equal to 1 in 2005 and multiply the result by the ICP price levels calculated above. The resulting series gives the price level of a particular sector in each country relative to the price of the same sector in the US in 2005.

Although the ICP study is especially built to provide accurate cross-country measures of price differences, it does have some well known limitations. For our purposes, the main objection is that expenditures are valued at the actual transaction prices paid by purchasers and hence may include delivery charges and any taxes payable (or subsidies received) on purchased products. This may be an issue if taxes/subsidies vary systematically with resource wealth. We recognize this fact, but our hands are tied for lack of better data. In the main body of the paper, we use a simple version of our model to show that to account for observed productivity differences, unrealistically large subsidies would be necessary. Notice also that this re-basing is not driving our results and we see similar productivity differences when value-added is left in constant US dollars.

**Sectoral value-added in International Dollars** We obtain one digit ISIC v.3 sectoral value-added data from the UN. UN data is given in constant 1990 USD prices and current prices.<sup>28</sup> First, we re-base the 1990 data to 2005 prices by calculating (for each sector) the ratio between the 2005 current and constant value-added. This gives us a relative sectoral price between 2005 and 1990:  $P_s^{2005}/P_s^{1990}$ . Multiplying the constant 1990 value-added series

<sup>28</sup> value-added by Economic Activity at constant 1990 prices, USD and value-added by Economic Activity at current prices, USD.

for each sector by this sector-specific price we obtain constant price sectoral value-added data in 2005 prices. Next, we need to convert the constant price (2005) sectoral value added data into one measured in international (or PPP) dollars. To do this, we divide constant (2005) price sectoral value-added data by the relative price levels,  $P_s^i/P_s^{PPP}$ , from expression 27. This converts sectoral value-added calculated in constant (2005) country specific prices into sectoral value-added calculated at international (2005) prices that are (in principle) invariant across countries and time. We recognize that these are imperfect price indices, however they are the best available, given data constraints. Finally, it is important to note that our empirical results do not - in any way - hinge on this procedure.

**Aggregate Capital** We follow Caselli (2005) and use the Penn World Tables (version 6.3) to construct estimates of aggregate capital stock. This is done using the perpetual inventory equation:

$$K_{t+1} = (1 - \delta)K_t + I_t, \quad (28)$$

where  $I_t$  is investment and  $\delta$  is the depreciation rate. Like Caselli (2005), we measure  $I_t$  from the PWT 6.3 as real aggregate investment in PPP.<sup>29</sup> As is standard, we compute the initial capital stock  $K_0$  as  $I_0/(g + \delta)$ , where  $I_0$  is the value of the above investment series in the first period that it is available, and  $g$  is the average geometric growth rate for the investment series in the first twenty years the data is available.<sup>30</sup> As is discussed in the literature - and by Caselli (2005) - the choice for initial capital stock is tenuous and stems from the assumption that an economy is on a balanced growth path of a Solow model (with a trend growth rate of  $g$ ) in the initial year. Finally, I follow Caselli (2005) and set the depreciation rate,  $\delta$ , to 0.06. The results prove not to be very sensitive to choices in either  $\delta$ ,  $g$  or initial capital stock.

The above process gives us sequences of capital stocks derived from PWT data. Notice however, that since we will be using UN PPP value-added data to calculate sectoral total factor productivity (and not PWT data), we want the capital values to be consistent with our UN total value-added data. As such, we use the PWT data to calculate the (unitless) capital-output ratio,  $k_t \equiv K_t/Y_t$  (where  $Y_t = RGDPL \cdot POP$  and  $K_t$  are both from the PWT data) and then use this ratio to construct a capital measure in terms of UN data:  $K_t = k_t \cdot VA_t$ , where  $VA_t$  is the UN measure of total value-added in 2005 international dollars, calculated previously.

**Sectoral Capital** We follow Caselli (2005) in estimating sectoral capital. First, assume that economies consist of five sectors: agriculture (A), mining and utilities (MU), manufacturing (M),

<sup>29</sup> So that  $I_t \equiv RGDPL \cdot POP \cdot KI$ , where  $RGDPL$  is real income per capita obtained with the Laspeyres method,  $POP$  is the population and  $KI$  is the investment share in real income.

<sup>30</sup> Caselli (2005) uses the growth rate between the first available year and 1970. We prefer our method, which should provide better estimates for countries whose investment data series start closer to 1970.

construction (C) and services (S). Then, assume that the production function of each sector,  $s$ , is of the form given in equation 2. If we also assume that the rates of return on capital are equalized across sectors (an arbitrage condition), then it is easy to show that the above functional forms implies that for any two sectors  $s$  and  $s'$ , the following holds:

$$\alpha_s \frac{P_s^D Y_s}{K_s} = \alpha_{s'} \frac{P_{s'}^D Y_{s'}}{K_{s'}}, \quad (29)$$

where  $P_s^D$  is the domestic producer price of sector  $s$  goods. As is emphasized by Caselli (2005), this price will generally differ from the PPP price and it is the price that a domestic investor will care about. Finally,  $P_s^D Y_s$  is sector  $s$ -es value-added (in domestic prices), calculated using UN current price data in local currency units. The above expression provides four distinct equations. Therefore, combining these with a capital market clearing condition:

$$\sum_s K_s = K, \quad (30)$$

where  $K_s$  is sector-specific capital and  $K$  is aggregate capital stock, we have a system of five equations in five unknowns from which we obtain an expression for sector specific capital stock,  $K_s$ , for each of the five sectors:

$$K_s = \left( \frac{\alpha_s P_s^D Y_s}{\sum_i \alpha_i P_i^D Y_i} \right) K. \quad (31)$$

Finally, to calculate the above expression we take labor shares,  $1 - \alpha_s$ , for each sector  $s$  from Table 6. Given these shares, we can use equation 31 for each sector and the aggregate capital stock (calculated previously) to obtain an estimate of sectoral capital.

**Aggregate Human Capital** We follow Caselli (2005) and Hall and Jones (1999) in constructing a measure of aggregate human capital. From the data set of Barro and Lee (2010) we obtain the average years of schooling,  $x$ , in the population over 25 years old. The schooling data is observed every five years, from 1950 up to (and including) 2010. Since  $x$ , moves slowly over time, we estimate the missing data by linear interpolation. This data is then turned into a measure of human capital,  $h$ , through the formula:

$$h = e^{\phi(x)}, \quad (32)$$

where  $x$  is the average years of schooling and the function  $\phi(x)$  is piecewise linear and defined as:

$$\phi(x) = \begin{cases} 0.134 \cdot s & \text{if } x \leq 4 \\ 0.134 \cdot 4 + 0.101 \cdot (x - 4) & \text{if } 4 < x \leq 8 \\ 0.134 \cdot 4 + 0.101 \cdot 4 + 0.068 \cdot (x - 8) & \text{if } 8 < x \end{cases} \quad (33)$$

The rationale for this functional form, as explained by Caselli (2005), is as follows:

Sector	Education Distribution							Ave. Y.	Ave. Years/Mfg
	<HS	HS	<C	C(A)	C(B)	M	D		
Agr	24.7	31.7	16.5	5.8	14.5	5.4	1.5	12.49	0.97
Constr.	21.4	39.8	20.4	6.3	9.0	2.3	0.6	12.18	0.94
Ser	7.8	24.3	21.3	9.3	23.1	9.7	4.6	14.22	1.10
Mfg	14.9	36.8	21.5	7.8	13.5	4.2	1.3	12.89	1.00
MU	13.0	33.3	22.6	8.8	15.4	5.1	1.7	13.18	1.02
Total	10.0	27.2	21.2	8.8	20.6	8.3	3.8	13.87	

Table 8: This table shows the distribution of education levels of workers in different ISIC sectors. The levels of education are: Less than a high school diploma (<HS); High school diploma or equivalent (HS); Some college, no degree (<C); Associate’s degree C(A); Bachelor’s degree C(B); Master’s degree (M) and Doctoral/professional degree (D). The table also shows the total implied average years of education by sector and relative to manufacturing. (Source: BLS)

Given our production function, perfect competition in factor and good markets implies that the wage of a worker with  $x$  years of education is proportional to his human capital. Since the wage-schooling relationship is widely thought to be log-linear, this calls for a log-linear relation between  $h$  and  $x$  as well, or something like  $h = e^{\phi_x x}$ , with  $\phi_x$  a constant. However, international data on education-wage profiles from Psacharopoulos (1994) suggests that in Sub-Saharan Africa (which has the lowest levels of education) the return to one extra year of education is about 13.4 percent, the World average is 10.1 percent, and the OECD average is 6.8 percent. Hall and Jones’s measure tries to reconcile the log-linearity at the country level with the concavity across countries.

**Estimating Education by Sector in the United States** Estimates for sectoral human capital,  $h_s$ , are very difficult to come by. As with aggregate human capital, these measures are often based on years of schooling in a particular sector - but this data is not readily available for most countries. When comparing agriculture and non-agriculture, Caselli (2005) infers the years of education in non-agriculture by assuming zero years of schooling in agriculture. Since we are interested in manufacturing versus non manufacturing data, we cannot follow this method. Instead, we base our sectoral educational estimates on US schooling data. In particular, using BLS data, we estimate the average years of schooling of those working in each ISIC one digit sector in the United States in 2008. The results are shown in Table 8. We then assume that the relative number of years of education between any two sectors remains constant (and the same as the US) across countries and time which allows us to infer sectoral education levels in all countries. To construct Table 8 from the BLS we obtain a distribution of occupations within each ISIC sector that specifies what fraction of employees within that sector work in a given

occupation.<sup>31</sup> We also obtain the economy-wide distribution of educational achievement for each occupation which describes what fraction of people in a particular occupation have achieved: (1) Less than a high school diploma; (2) a High school diploma or equivalent; (3) Some college, no degree; (4) an Associate's degree; (5) a Bachelor's degree; (6) a Master's degree and (7) a Doctoral/professional degree. Given these distributions, we can then calculate the distribution of educational levels within each ISIC sector.<sup>32</sup> Finally, assuming that education levels (1) through (7) take 8, 12, 14, 14, 16, 18 and 21 years respectively to achieve, allows us to calculate the average number of years of education of people working within each ISIC sector.

**Sectoral Human Capital** To calculate sectoral human capital we assume that the relative number of years of education between any two sectors remains constant (and the same as the US) across countries and time. More specifically, we define  $\eta_s^{US}$  as the ratio of average years of schooling in sector  $s$  relative to the years of schooling in the manufacturing in 2008 in the US (from Table 8):

$$\eta_s^{US} = x_s^{US} / x_m^{US}. \quad (34)$$

We then assume that for country  $i$ , the average years of schooling in sector  $s$ ,  $x_s^i$ , is related to the number of years of schooling in manufacturing in that country by:

$$x_s^i = \eta_s^{US} x_m^i. \quad (35)$$

We are thus assuming that the relative number of years of education between any two sectors remains constant (and the same as the US) across countries and time. Finally, education must also satisfy the following aggregation identity for each country:

$$\sum_s l_s^i x_s^i = x^i, \quad (36)$$

where  $l_s^i$  is the employment share of sector  $s$  in country  $i$  (so that  $\sum_s l_s^i = 1$ ) and  $x^i$  is the average years of schooling per worker in the entire economy. Given employment shares, the aggregate years of schooling and  $\eta_s^{US}$ , the above expressions yield five equations in five unknowns, which can be solved for years of schooling in each sector and country,  $x_s^i$ . For each country, we can then relate the years of schooling in each sector to sectoral human capital through the 'standard' Mincerian returns formula in equation 33.

<sup>31</sup> Occupations are classified by major - two digit - 2010 Standard Occupational Classification (SOC).

<sup>32</sup> For example, suppose  $z_A$  is a vector that contains the distribution of occupations within agriculture and  $w^{HS}$  is the distribution of those workers who have achieved at most a high school degree across all occupations. Then the dot product of the two vectors,  $z_A \cdot w^{HS}$ , is the fraction of agricultural workers who have achieved at most a high school degree.

**Public Sector Employment** Public sector employment data is from the ILO which “covers all employment of (the) general government sector as defined in System of National Accounts 1993 plus employment of publicly owned enterprises and companies, resident and operating at central, state (or regional) and local levels of government. It covers all persons employed directly by those institutions, without regard for the particular type of employment contract.” A limited subset of the public employment data is provided at the ISIC one sector level and in that (very limited) subset, public employment is overwhelmingly in the non-manufacturing sector. As such in the baseline experiment of this paper - in order to maintain as large a sample of data as possible - I shall assume that all government employment belongs entirely to the non-manufacturing sector.

## References

- Barro, Robert and Jong-Wha Lee**, “A New Data Set of Educational Attainment in the World, 1950-2010,” *NBER Working Paper No.15902*, 2010.
- Caselli, Francesco**, *Handbook of Economic Growth, Volume 1A.*, Elsevier B.V.,
- Collier, Paul and Anke Hoeffler**, “Resource Rents, Governance, and Conflict,” *Journal of Conflict Resolution*, 2005, *49(4)*, 625–633.
- **and Benedikt Goderis**, “Commodity Prices, Growth, and the Natural Resource Curse: Reconciling a Conundrum,” *CSAE Working Paper*, 2009, *2007-15*.
- Corden, W Max and J Peter Neary**, “Booming Sector and De-Industrialisation in a Small Open Economy,” *The Economic Journal*, 1982, *92*, 825–48.
- Dekle, Robert and Guillaume Vandembroucke**, “A Quantitative Analysis of Chinas Structural Transformation,” *Federal Reserve Bank of San Francisco Working Paper*, 2011.
- Duarte, M. and D. Restuccia**, “The Role of the Structural Transformation in Aggregate Productivity,” *Quarterly Journal of Economics*, 2010, *125(1)*, 129–173.
- Eaton, J. and S. Kortum**, “Technology, Trade and Growth: A Unified Framework,” *European Economic Review*, 2001, *45*, 742–755.
- Gollin, Douglas, Stephen Parente, and Richard Rogerson**, “The Role of Agriculture in Development,” *American Economic Review*, 2002, *92(2)*, 160–164.
- Haan, Laurens De and Ana Ferreira**, *Extreme Value Theory: An Introduction*, Birkhauser (Springer), 2006.
- Hall, Robert E. and Charles Jones**, “Why Do Some Countries Produce So Much More Output Per Worker Than Others?,” *The Quarterly Journal of Economics*, 1999, *114(1)*, 83–116.
- Heathcote, J., F. Perri, and G. Violante**, “Unequal we Stand: An Empirical Analysis of Economic Inequality in the United States, 1967-2006,” *Review of Economic Dynamics*, 2009, *Special Issue*.
- Heckman, James J. and Guilherme Sedlacek**, “Heterogeneity, Aggregation, and Market Wage Functions: An Empirical Model of Self-Selection in the Labor Market,” *The Journal of Political Economy*, 1985, *93(6)*, 1077–1125.

- ILO**, “Economically Active Population 1950-2010: ILO Database on Estimates and Projections of the Economically Active Population (5th edition),” 2003.
- Kuralbayeva, Karlygash and Radoslaw Stefanski**, “Windfalls, Structural Transformation and Specialization,” *Journal of International Economics*, 2013, 90, 273–301.
- Lagakos, D. and M . Waugh**, “Selection, Agriculture, and Cross-Country Productivity Differences,” *American Economic Review (forthcoming)*, 2012.
- Lucas, Robert**, “On the Size Distribution of Business Firms,” *The Bell Journal of Economics*, 1978, 9, 508–523.
- Matsuyama, Kiminori**, “Agricultural Productivity, Comparative Advantage and Economic Growth,” *Journal of Economic Theory*, 1992, 58, 317–334.
- Michaels, Guy**, “The long term consequences of resource-based specialisation,” *Economic Journal*, 2011, 121(551), 31–57. Mimeo, London School of Economics.
- Neary, J. Peter**, “Short-Run Capital Specificity and the Pure Theory of International Trade,” *Economic Journal*, 1978, 88, 488–510.
- Psacharopoulos, G.**, “Returns to investment in education: A global update,” *World Development*, 1994, 22(9), 1325–1343.
- Restuccia, Diego, Dennis Tao Yang, and Xiaodong Zhu**, “Agriculture and Aggregate Productivity: A Quantitative Cross-Country Analysis,” *Journal of Monetary Economics*, 2008, 55(2), 234–250.
- Robinson, James, Ragnar Torvik, and Thierry Verdier**, “Political foundations of the resource curse,” *Journal of Development Economics*, 2006, 79, 447–468.
- Rogerson, R.**, “Structural Transformation and the Deterioration of European Labor Market,” *Journal of Political Economy*, 2008, 116(2), 235–259.
- Roy, A. D.**, “Some Thoughts on the Distribution of Earnings,” *Oxford Economic Papers*, 1951, 3(2), 135–146.
- Sachs, Jeffrey D. and Andrew M. Warner**, “The Curse of Natural Resources,” *European Economic Review*, 2001, 45, 827–838.
- UN**, “United Nations National Accounts Statistics Database,” 2008.
- van der Ploeg, Frederick**, “Natural Resources: Curse or Blessing?,” *Cesifo Working Paper*, 2010, No. 3125.

**WDI**, “World Development Indicators,” 2007.

**WEO**, “World Energy Outlook,” 2011.

**Yi, Kei-Mu and Jing Zhang**, “Structural Change in an Open Economy,” *Working papers in International Economics, University of Michigan*, 2010, No. 595.