

Evaluating Explanations for Stagnation*

Elizabeth M. Caucutt

University of Western Ontario
ecaucutt@uwo.ca

Krishna B. Kumar

USC Marshall School of Business
kbkumar@usc.edu

December 2004

Abstract

In this paper, we evaluate four explanations for economic stagnation that have been proposed in the literature: coordination failures, ineffective mix of occupational choices, insufficient human capital accumulation, and politico-economic considerations. We calibrate models that embody these explanations in the context of the stagnant economies of sub-Saharan Africa. The methodology of calibration is ideally suited for this evaluation, given the paucity of high-quality data, the high degree of model nonlinearity, and the need for conducting counterfactual policy experiments. In addition to studying how closely and robustly these models capture the African situation, we examine the quantitative aspects of their policy implications. We find that calibrations that yield multiple equilibria – one prosperity and the other stagnation – are not particularly robust. This tempers optimism about the efficacy of one-shot or temporary development policies suggested by models with multiplicity. However, the calibrated models indicate that small policy interventions are sufficient to trigger development in stagnant economies.

Keywords: Coordination failure, Occupational choice, Human capital accumulation, Political economy, Economic Development, Calibration.

JEL Classification: O100, O110, E600

*We are grateful to Daron Acemoglu, Michele Boldrin, Oded Galor, Doug Joines, Larry Jones, Pete Klenow, Nancy Stokey, Chris Udry, and participants at various seminars and conferences for their comments and suggestions.

1 Introduction

A substantial branch of the literature on economic growth and development is devoted to understanding stagnation, a condition in which economies are locked into low output and income. Understanding the causes of stagnation and policies to overcome it has immense implications for human welfare. Most papers in this tradition develop theoretical models to highlight a particular economic force, and provide conditions under which stagnation, or a “trap” as it is often referred to, results. Policy implications are mentioned, but typically not analyzed or quantified in detail.

How well do these models explain the economic stagnation seen in the data? What are the quantitative implications of policies they suggest? Since the reasons for stagnation are likely to be as multifarious as engines of growth, are there any policy lessons to be learned by considering these models collectively instead of one at a time? These are a few of the questions we address in this paper, by applying the methodology of calibration to selected models. The poorest countries in sub-Saharan Africa (sSA), in which per capita income and output have been low and stagnant during the last three to four decades, provide a natural context for such an evaluation. To the best of our knowledge, ours is the first attempt at a quantitative and comparative evaluation of models of stagnation.¹

Calibration is ideally suited for the study of stagnant environments, where the scarcity of high-quality data makes detailed econometric analysis, especially at the macroeconomic level, difficult.² Calibration also readily lends itself to analyzing counterfactual policy experiments that can pry an economy out of stagnation.

We evaluate four explanations for stagnation found in the literature: 1) Unresolved coordination problems in the presence of increasing returns, 2) Occupational choices detrimental to development arising from imperfect capital markets, 3) Insufficient human capital accumulation, also driven by capital market imperfections, and 4) Political economy considerations that lock an economy into a low-performing state.

We choose models that are representative of each explanation for our calibration exercise. While other explanations and models could be found, the ones chosen do provide enough diverse dimensions for conducting a comparative quantitative analysis of the problem of economic stagnation. We distinguish between parameters that are “structural” in the sense that they are expected to hold everywhere, and those particular to sSA that cause stagnation

¹See Azariadis (1996), and Bowles, Durlauf, and Hoff (2003) for state-of-the-art surveys on models of poverty traps. For detailed evidence on the stagnancy of sub-Saharan Africa see Acemoglu, Johnson, and Robinson (2002) and Caucutt and Kumar (2004).

²For examples of econometric work, see Durlauf and Johnson (1995), who find multiple regimes in cross-country dynamics, Quah (1996), who studies distribution dynamics, and McKenzie and Woodruff (2002), who find little evidence for production non-convexities as a source of poverty traps among Mexican microenterprises.

in a given model. For structural parameters we use the more readily available data from developed countries. For particular parameters, we use data from sSA, from whichever country and source it is available, and rely on ranges of estimates where needed.

Once we evaluate the robustness of a model in producing the stagnant outcome, we design and implement policy experiments that are appropriate to the model. We quantify each policy in terms of tax rates, cost of subsidies as a fraction of GDP, or welfare gain in terms of equivalent variation so that we can compare policies across models. Mauritius, a successful economy in sSA, often serves as an empirical anchor against which we assess a model's policy recommendations.

To study coordination problems, we calibrate the “Big Push” models of Murphy, Shleifer, and Vishny (1989), which feature expectations-driven multiple equilibria. Each sector in the economy is willing to incur a fixed cost and implement a labor saving technology if it expects all other sectors to do so, but not otherwise. We can find reasonable parameters for which this multiplicity results. However, for this and most other models we study, multiplicity is not particularly robust to changes in parameters in the direction of greater empirical plausibility. Conditional on multiplicity, a fairly low rate of one-time investment subsidy, 4 to 7% for most parametrizations, is enough to avoid stagnation. The drive toward industrialization by Mauritius in the 70s, using investment tax subsidies and other incentives to foster export processing zones, provides empirical support for this type of policy intervention.

We calibrate the Banerjee and Newman (1993) model to study the role of occupational choice in stagnation. In this model, imperfect enforcement in the capital market motivates collateral-based lending for project financing. Based on the level of their initial wealth, agents choose to be workers, self-employed, or entrepreneurs. If the starting ratio of workers to entrepreneurs is low, the dynamics is characterized by high wages and a prosperous steady state will be reached. However, if this ratio starts off high, the wage remains low, and the economy is trapped in an absorbing, subsistence state. A restricted set of parameters yields this multiplicity, but we are able to map initial wealth distributions of Tanzania and Mauritius to the model, and demonstrate how a “bad” initial distribution could have led Tanzania toward stagnation and a “good” one led Mauritius to prosperity. The one-time redistribution needed to change the distribution from bad to good is about 4% of total wealth.

We develop and calibrate our own model, which is in the spirit of Galor and Zeira (1993), to study the human capital explanation. In this heterogeneous-agent model, developed with the explicit intent of calibration, high costs of education relative to low income and skill premium cause the economy to stagnate in a low steady state with minimal educational attainment. While the model is theoretically capable of yielding multiple steady states, the benchmark calibration yields only a single stagnant steady state. A continual tax and in-kind subsidy that effectively redistributes resources from poor households with lower ability children to those with higher ability children can pry the economy out of stagnation, freeing

it from dependence on foreign aid. We find that a GDP share of education of 3.2% is required in the calibrated model to produce a Mauritius-like outcome, close to the actual expenditure share seen in data.

Finally, we consider the more recent politico-economic literature on stagnation. Unlike the previous models, there is no multiplicity of equilibria here.³ We first calibrate the model of Bourguignon and Verdier (2000), who assume that only an educated-wealthy minority (oligarchy) has political power. There are external benefits to education, which provide the oligarchs with an incentive to subsidize the education of the poor. However, the newly educated poor may gain political control, and impose costly redistributive taxes. Under certain conditions, democratization, which in the model is equivalent to education of the poor, occurs. We calibrate the model to Tanzanian data and find that the no-democratization condition is met, though a higher than plausible education externality is required. An education subsidy scheme in which costs are shared between the local wealthy minority (5.7% of their lifetime income) and foreign aid (8.7% of GDP for seven years) can move Tanzania to a fully educated democracy.

Acemoglu and Robinson (2002), consider a “ruler” who trades off the benefit of adopting a new technology and getting a percentage of the increased output as taxes, against the cost of being replaced due to the increased “turbulence” that results from the adoption. Rulers with a firm grip on power, as well as those facing a high probability of replacement, innovate, while those in the middle block technology adoption. We find that parameters drawn from innovating countries can be consistently used to explain politico-economic stagnation in sSA. However, the model is not amenable to quantifying the costs of policy reforms such as increasing democratization or lowering political rents.

The models we evaluate are stylized, each abstracting along several dimensions in order to focus on its main channel.⁴ Moreover, comprehensive, good-quality data is rarely available for a particular sSA country. Both these considerations necessitate a flexible approach toward calibration. Sometimes, this involves finding any set of parameters that can produce the stagnation outcome of the model and then evaluating the empirical validity of the parameters, rather than starting with parameters that seem *a priori* reasonable. In order to preserve other authors’ original intent, we do not modify their models. Our aim is not to merely survey these models; our calibration and policy experiments are original additions that subject these models to the rigor of quantitative analysis. None of these models attempts to explain all the income differences seen in the data. Therefore, we confine our quantitative analysis to the stagnation seen in sSA.

³Azariadis and Drazen (1990), Becker, Murphy, and Tamura (1990), and Durlauf (1993) are a few of the other models that feature multiplicity.

⁴As Murphy, Sleifer, and Vishny (1989) note, “... because all our models are highly stylized and capture what we can only hope to be one aspect of reality, policies suggested by these models should be interpreted with caution.” (p. 1006)

What answers can we provide to the questions that motivated this study? First, it is possible to find parametrizations for all models – some empirically more reasonable than others – that are consistent with stagnation in sSA. But across the models, we find that calibrations that yield multiple equilibria – one prosperity and the other stagnation – are not particularly robust. When there are multiple steady states, a one-time policy intervention can alter the initial condition and steer the economy toward the high development steady state instead of stagnation. On the other hand, when there is a unique low development steady state, the policy change has to be permanent.⁵ Given the difficulty of obtaining multiple equilibria, we see the need for caution in advocating one-shot or temporary policies.

Second, our analysis of policy reforms shows that the costs of implementing the policies suggested by the models are not very high. This is true even in our human capital model, where we study only permanent policies.

Third, even though the models focus on different channels, we find that a more complete picture of the causes of stagnation and policy reforms emerges when we consider them together. For instance, while our model suggests foreign aid is not required for development, Bourguignon and Verdier’s (2000) political model shows that foreign aid can act as a seed in encouraging human capital investments by the local wealthy minority. In general, the political economy models provide insights into why the seemingly small interventions suggested by other models might be difficult to implement. Implementing democratic reforms and providing foreign assistance in alignment with the interests of a local oligarchy that can block reform are inherently difficult to do. Considering the models together also allows us to identify recurring explanatory factors – the initial income distribution, human capital, and capital market imperfections – which would be leading candidates for inclusion in a more comprehensive model of stagnation.

Beyond this, our exercise naturally allows us to identify the relative quantitative strengths and weaknesses of each model. For instance, the parameters of the political economy models are harder to calibrate than the technological variables of the other models. However, these models are the most parsimonious with the fewest number of parameters to calibrate. The calibrations of the coordination failure and occupational choice explanations need to rely less on sSA-specific parameters, which is useful given the erratic availability of data; but the other models that do rely on specific parameters have the advantage of explaining a particular rather than a generic situation of stagnation. We defer a more detailed comparative evaluation until Section 6, and first present the individual analyses. Sections 2 through 5 consider, respectively, the explanations of coordination failure, occupational choice, human

⁵As Banerjee and Newman (1993) note, “Under the guidance of the linear model, which usually displays global stability, one is led to conclude that continual redistributive taxation, with the distortion it often entails, is required for achieving equity. The nonlinear model, by contrast, raises the possibility that one-time redistributions may have permanent effects, thereby alleviating the need for distortionary policy.” (p. 296)

capital accumulation, and the political economy. For each, we present a brief summary of the model, the calibration strategy, the potential of the calibrated model to explain stagnation, and the outcome of policy experiments. Section 7 concludes.

2 Coordination Failure

We consider the work of Murphy, Shleifer, and Vishny (1989) to analyze coordination failure. Here, a firm’s investment exerts a pecuniary externality on other firms by increasing the market size or decreasing infrastructure costs. Since individual firms do not take this effect into account, there could be a coordination failure which causes stagnation. Coordination of investment across sectors could give the economy a “Big Push” and move it to the good equilibrium; simultaneous industrialization could be self-sustaining even if a sector cannot afford to industrialize on its own.

2.1 Model

Murphy, Shleifer, and Vishny (MSV) first consider a unit interval of goods with the utility function, $\int_0^1 \ln x(q) dq$, which implies equal expenditure shares. There are L units of labor, with wage being the numeraire. Each sector has a competitive fringe, which converts labor to output one for one, and a potential monopolist with an increasing returns to scale technology, each unit of labor yielding $\alpha > 1$ units of output. For a firm to acquire the increasing returns technology (become “industrialized”) and gain monopoly over an entire sector, it has to incur a fixed cost of F units of labor. Since the firm faces the entire demand curve for the good, given income y , the firm’s profit is $\pi = ay$, where $a \equiv (1 - 1/\alpha)$ is the markup. If n sectors industrialize, aggregate profits are $\Pi = n\pi$. These are repatriated to the households, implying an income of $y = \Pi + L$. Without any industrialization, income is L . The income increases with the degree of industrialization, n ; an industrializing sector gives profits back to consumers who spend it on *all* goods and raise the profits of *all* industrialized firms. This basic setup gives only one equilibrium – stagnation or industrialization – depending on the parameters. If it is unprofitable for one firm to industrialize when its income is only L , and if it industrializes anyway, it reduces aggregate income making it more unprofitable for all other firms to industrialize. MSV then present three extensions to ensure a firm that engages in an unprofitable investment can still benefit other sectors, making it likely they find investment profitable. This yields multiple equilibria and the possibility of a Big Push.

The first extension assumes that to attract workers away from CRS farm work to IRS manufacturing, firms have to pay a premium, since working in factories entails a disutility of v . Given a farm wage of one, the factory wage is $1 + v$. The condition for no industrialization (stagnation) to occur is $L(1 - (1 + v)/\alpha) - F(1 + v) < 0$. If a firm expects no other firm to industrialize, and therefore aggregate income to be L , it does not incur the fixed cost of F units of factory labor. The condition for all firms to expect a high level of income and

sales from simultaneous industrialization and be willing to incur the fixed cost is $\alpha(L - F) - L(1 + v) > 0$. If both conditions are satisfied, both equilibria are possible. It is convenient to write the condition that parameters need to satisfy for multiplicity as

$$(1 + v) < \alpha(1 - F/L) < (1 + v) + \alpha v F/L. \quad (1)$$

The second extension is a two-period model of investment, with the extended utility specification $\left[\int_0^1 x_1^\gamma(q) dq\right]^{\frac{\theta}{\gamma}} + \beta \left[\int_0^1 x_2^\gamma(q) dq\right]^{\frac{\theta}{\gamma}}$; the intertemporal elasticity of $1/(1 - \theta)$ and elasticity of substitution across goods is $1/(1 - \gamma)$. The discount factor is β . In the first period, only the CRS technology is available. This is also available in the second period; however, a potential monopolist can invest F units of labor in the first period to acquire the IRS technology in the second period. The profit for such a monopolist is given by $\pi = (1/(1 + r)) ay_2 - F$, where r is the interest rate, y_2 the second period income, and a is the markup defined earlier. The condition for no sector to industrialize is $(1/(1 + r)) aL - F < 0$. The demand firms expect to obtain in second period is too low for them to break even on their investments, and the realized income is indeed low. The income of L in each period is consistent with the interest factor $(1/(1 + r)) = \beta$. The condition for an industrialized equilibrium is $(1/(1 + r)) a\alpha L - F > 0$, where the interest factor consistent with a first period income of $(L - F)$ and a second period income of αL is $(1/(1 + r)) = \beta(\alpha L/(L - F))^{\theta - 1}$. The increase in investment demand by the firms increases the interest rate, decreasing the discount factor a firm uses to assess profitability. The effect of increased income from monopoly profits (repatriated to consumers) has to dominate this decrease in the discount factor. Again for some parameter values both conditions are met. The condition for multiplicity is

$$\frac{1}{\alpha^\theta \left(1 - \frac{F}{L}\right)^{1 - \theta}} < \beta a < F/L, \quad (2)$$

which uses the above-mentioned interest factors.

The third extension considers an investment in infrastructure, say a railroad. The $\theta = 1$, $\gamma = 0$, version of the above utility is used. Though MSV ignore β by setting it to one, we retain it to facilitate realistic calibration. CRS technologies can be set up anywhere and don't use the railroad. IRS technologies are location specific and need the railroad to sell their products. A fraction n of the sectors need a first-period fixed cost of F_1 units of labor to industrialize while the remaining $(1 - n)$ need fixed cost $F_2 > F_1$. It costs R units of labor to build the railroad in first period and marginal cost of its use is zero. The type of the firm is private information and the monopolist cannot price discriminate. It is assumed that even if all type 1 firms industrialize, the surplus generated will not cover R ; both types of firm must industrialize.

There are two considerations – whether the railroad is built even if it is efficient, and whether multiplicity can exist even if the railroad is built. The condition for an equilibrium in which the railroad is built and all sectors industrialize is $(1/(1 + r)) a\alpha L - F_2 > R$.

Given the inability to price discriminate, the railroad company extracts all the surplus of high-cost firms and extracts the same from low-cost firms, leaving them with a positive surplus. With $\theta = 1$, there is no interest rate effect and $(1/(1+r)) = \beta$. Even when railroad building is efficient, that is, when $(1/(1+r))a\alpha L - nF_1 - (1-n)F_2 > R$, if the stronger industrialization condition is not satisfied, the railroad will not be built. The condition for no industrialization is $(1/(1+r))aL - F_1 < 0$, where the interest factor is β . The condition for multiplicity is, therefore

$$(F_2/L + R/L)/\alpha < \beta a < F_1/L. \quad (3)$$

If this condition holds, the uncertainty concerning equilibrium selection might cause the railroad to not be built, since the railroad will be profitable if the economy industrializes but incur a large loss if no industrialization occurs. This, in addition to the inability to price discriminate, might warrant subsidization of railroad construction. Additionally, coordination of investments might be required to avoid multiplicity of equilibria.

2.2 Calibration

We realize the need to be flexible in calibrating stylized models written with the aim of highlighting particular forces of economic development analytically. Therefore, for this and other models, we sometimes resort to parameters that “work”, and then assess if they are empirically reasonable, rather than follow the usual strategy of fixing some of the parameters *a priori* using independent evidence, and calibrating others to match empirical targets.

The parameters that are common to all three models are α , the degree of increasing returns, and F/L , the normalized cost of adopting the increasing-returns technology. Hall (1988) presents estimates of the markup ratio (price to marginal cost) in the US economy, which corresponds to the α of the MSV model. The estimates for one-digit industries range from 1.864 for services to 3.791 for trade. Hall (1990) presents direct evidence on the IRS parameter, which ranges from 1.08 in services to 10.03 for transportation. We find that a value of $\alpha = 3$, which is roughly the value for nondurables in both estimates, works for all three models. This might seem like a high value, especially in light of the highest value of 1.72 reported in Basu and Fernald (1997) for the entire private economy. However, the ratio of the income between the industrialized and non-industrialized economies at the end of the second period in the two-period models is also α , and from this viewpoint a value of 3 does not appear to be too high.

In the context of the MSV model, the quantity F/L can be interpreted as the fraction of either labor or resources devoted to technology adoption. We find the share of skilled labor in total costs a proxy that works. Greenwood and Yorukoglu (1997) view adoption in this fashion and turn to data from Bartel and Lichtenberg (1987) for empirical support. The data in Bartel and Lichtenberg indicates that the ratio of earnings of those with 13 or more years of education to those with less was fairly stable at 0.6 in US manufacturing from the 60s

through the 80s. This implies a skill share in total labor costs of 0.375. This is higher than other candidate proxies – the actual employment share of educated workers, which ranges from .158 to .271 during that period, and the 12% of the population who are entrepreneurs (De Nardi and Cagetti (2003)).

For the factory premium model, we use the rural-urban wage gap to proxy for the factory disutility v . The USDA’s Economic research service reports rural wages of about 70% of urban wages, which implies $v = 0.4286$. This is comparable to the value of 0.376 derived from BLS weekly earnings ratio in the goods-producing industry to the private service-providing industry.⁶

For the investment model, we assume a base annual real interest rate of 8%. This implies an annual β of 0.9259.⁷ This annual value has to be compounded over a long enough gestation period that is typical of large-scale industrial projects. Given the varying gestation periods observed for different industries, we study a range of values for the compounded β : 0.5401 for 8-year compounding, 0.4631 for 10 years, and 0.3151 for 15 years. Whether the industrialization condition is met or not depends strongly on the value used for θ , which controls the effect of deferred consumption on the interest rate. The easiest to consider is $\theta = 1$, which implies infinite substitutability across periods. An alternate interpretation is a small open economy in which there are no interest rate effects of increased investment. Other values we use for θ range from 0.35 to 0.75, which yield elasticity of substitutions of 1.5 to 4. While this elasticity is not far from the value of 1 often used in the calibration of macroeconomic models, it is much larger than the 0.2 to 0.4 figure reported by Patterson and Pesaran (1992) for the US and UK, which imply a θ of -4 to -1.5.⁸ Any value of θ less than 0.35 causes the interest rate effect to dominate and makes industrialization impossible.⁹

For the railroad cost, R/L , in the infrastructure model, we use the information from the *World Development Report 1994*, that public infrastructure investment in developing countries ranges from 2 to 8%, with an average of 4%. These figures are in the ballpark of the US infrastructure spending, which was between 2.5% and 3% of GDP during 1956-1991.¹⁰ We use the figures of 3% and 8% to cover a broad enough range. This model extension also

⁶The values chosen for α and v satisfy the MSV condition of $\alpha - 1 > v$, for the increasing returns to be sufficiently high to warrant the higher factory wages.

⁷As will be seen, what matters is the annual β compounded by a gestation period. Different values of the annual β and years are therefore compatible with the final values used.

⁸Güvenen (2003) argues that a higher elasticity would result when wealth data, rather than consumption data, is used. These higher estimates are connected to aggregate investment and output, and seem more relevant for our purposes.

⁹MSV assume $\alpha < 1/\gamma$ to ensure a sufficiently inelastic demand. For our chosen value of α , we need $\gamma < 1/3$, which implies an elasticity of substitution among goods of less than 1.5. For instance, the Cobb-Douglas aggregator $\int_0^1 \ln x_i(q) dq$, $i = 1, 2$, satisfies this constraint.

¹⁰See the Congressional Budget Office’s, *Trends in Public Infrastructural Spending*, 1999.

requires that entry costs be broken into low and high costs. In line with our calibration of F/L , we use for F_1/L and F_2/L , the lower and upper ends of the range of labor cost share of highly educated workers as reported in Bartel and Lichtenberg (1987): 0.307 for Wood Containers, and 0.433 for Electronic Components.¹¹

Table 1: Parameters for the Murphy-Shleifer-Vishny Models

Parameter	Value	Comment
α	3	IRS parameter for nondurables; Hall (1988, 1990)
F/L	0.375	Skill share in total labor costs; Bartel and Lichtenberg (1987)
v	0.4286	From rural-urban wage gap; USDA
β	0.3151 - 0.5401	Annual value of 0.9259 compounded over 8 to 15 years
θ	0.35 - 0.75	“Closed” economy; intertemporal elasticity of 1.5 to 4
	1	“Open” economy; infinite intertemporal elasticity
R/L	0.03 - 0.08	Public infrastructure investment; <i>World Development Report 1994</i>
F_1/L	0.307	Upper end of skill share in labor costs; Bartel and Lichtenberg (1987)
F_2/L	0.433	Lower end of skill share in labor costs; Bartel and Lichtenberg (1987)

Table 1 presents the parameters we use.¹² We have assumed that these parameters are “structural” and hold across all economies; this is easier to defend if we take the view that all parameters are purely technological. This allows us to assume, in the spirit of MSV, that some economies stagnate purely on account of bad expectations, and focus on Big Push policies to undo these expectations. However, we later discuss the possibility of identifying the fixed costs with business entry costs, which are much higher in sSA than in OECD countries and even the more prosperous countries within sSA.

2.3 What the Calibrated Models Explain

The benchmark values satisfy the multiplicity condition in the factory premium model, (1), by yielding: $1.4286 < 1.875 < 1.9108$. High enough values for v , F/L , and α are needed for the condition to hold. If the alternate value of 0.376 is used for v , the no-industrialization condition will cease to hold. This also happens if the share of entrepreneurs in the population, 0.12, is used for F/L . If a lower value of α , say 2, is used, the degree of increasing returns is not enough to warrant the fixed costs and only the no-industrialization condition is satisfied.

In the open economy version of the two-period investment model (with $\theta = 1$), the condition for multiplicity, (2), is satisfied for values of β involving 8 or more years of project gestation period. For instance, with 8-year compounding, the condition is satisfied as:

¹¹Another interpretation of differing fixed costs could be that some firms are more *efficient* than others at adopting similar technologies. However, given greater data availability, we have chosen the interpretation that differing industry-specific technologies are the source of different fixed costs.

¹²Recall that the measure of low-cost firms, n , enters the efficiency condition. Our benchmark parameters satisfy the stricter industrialization condition. Therefore, we do not need a value for n .

0.125 < 0.3602 < 0.375. Lower values for F/L will necessitate compounding by longer periods. The condition for no industrialization will cease to hold for a high β , low F/L combination – the discounted gains from industrialization exceed the cost. Greater variation in α can be tolerated for this model; for instance a value of 1.5 will work.

The closed economy version of the model is highly sensitive to the value assumed for θ , and also for β . The following combinations of parameters yield multiplicity when the benchmark value of F/L is used: $\theta = 0.35 - 0.5$ and $\beta = 0.5401$ (8 years), $\theta = 0.425 - 0.45$ and $\beta = 0.4631$ (10 years), and $\theta = 0.675 - 0.75$ and $\beta = 0.3151$ (15 years). Any value of θ lower than those specified will decrease the intertemporal elasticity of substitution enough to cause the interest rate effect to dominate; only the no-industrialization condition will be left standing. As in the open economy version, a high β , and a low F/L will break the no industrialization condition.

Finally, the infrastructure model satisfies the multiplicity condition (3) for a 11-year β of 0.4288. With an R/L of 0.03, the condition is satisfied as $0.1543 < 0.2858 < 0.307$, and also for an R/L of 0.08, with the first number now becoming 0.171. Since the relevant fixed cost is now lower (as $F_1/L < F/L$), any higher value for β will make the fixed investment attractive enough to break the no industrialization condition. Lower values of α , say 2.5, will help in this regard, but tighten the condition needed for industrialization.

In summary, it is possible to find parameters for which the multiplicity conditions hold for all three models of MSV, which warrants the study of policies involving a Big Push. However, the open economy version of the two-period investment model is the most robust to changes in parameters. A high enough value for fixed costs is crucial in all three models to satisfy the no-industrialization condition, which, given the interest in stagnation, is the more important of the two conditions.

2.4 Policy Experiments

The MSV models identify conditions under which a given set of parameters satisfy both industrialization and stagnation. However, they do not take a stance on equilibrium selection. Therefore, we assume that extrinsic conditions resulted in the selection of the stagnant equilibrium in sub-Saharan Africa (sSA). The task in this subsection then is to identify policies that would give a Big Push to the economy to pry it out of stagnation, and explore whether such policies have worked in the region. MSV mention the policies of investment subsidies and coordination, but do not explicitly analyze them. However, it is fairly straightforward to derive the minimum rate of investment subsidy required in each of the three models, to break the stagnation condition and spur industrialization. The aim would be to reduce the effective cost of fixed investment F by enough, say to $(1 - s)F$, such that even if a potential monopolist does not expect other sectors to industrialize, and therefore expects an aggregate income of only L , he would individually find it profitable to industrialize. We assume that

the cost of funding these subsidies, sF , is met by taxing income.¹³

First consider the factory premium model. We can convert the stagnation condition $L(1 - (1 + v)/\alpha) - F(1 + v) < 0$ into an industrialization condition by writing

$$(L - sF)(1 - (1 + v)/\alpha) - (1 - s)F(1 + v) > 0.$$

For the parameter values assumed in Table 1, this implies a minimum investment subsidy rate of 3.5%. As a fraction of the stagnant income L , the income tax payments, sF , are 1.3%.

The condition for stagnation in the two-period investment model is $(1/(1 + r))aL - F < 0$. However, we cannot set the interest rate factor to β , if we expect to fund investment subsidies from taxes on first period income. It would have to be consistent with the consumption of $L - sF$ and L in the two periods. Therefore, we write the condition for the required subsidy as

$$\beta((L - sF)/L)^{1-\theta} aL - (1 - s)F > 0.$$

Financing investment subsidies by taxing first period consumption automatically accomplishes the MSV recommendation of “discouraging current consumption.” The $\theta = 1$ case is again the easiest to consider. With a β computed for 8 years, the minimum subsidy rate required is about 4%; as a fraction of the stagnant income, income tax payments are 1.5%. These rates are highly sensitive to β . If a 10-year compounding is used, these rates increase to 17.3% and 6.5% respectively.

When $\theta < 1$, the first term in the above condition also decreases with s , reflecting the increased interest rate needed to induce consumers to postpone consumption. However, it declines less steeply in s than the second term, and a unique minimum subsidy rate, $0 < s < 1$, can be found. As θ (the intertemporal substitutability) decreases, the minimum subsidy increases; the monopolist discounts more due to the increased interest rate, and requires higher subsidies. When a 8-year β is used, the subsidy rate is 4.8% when θ is 0.5, and 5.2% when θ is 0.35; the corresponding income tax rates are 1.8% and 1.9%. As in the $\theta = 1$ case, these rates are highly sensitive to β . With a 10-year β , and θ set to 0.45, the minimum subsidy rate required is 20.9% and the income tax rate is 7.8%.

In the infrastructure model, in which θ has already been set to 1, and therefore the relevant discount factor for the monopolist is just β , the condition to overcome stagnation is simply

$$\beta aL - (1 - s)F_1 > 0.$$

Using the parameters of the previous subsection, we get a minimum subsidy rate of 6.9%, which translates to an income tax rate of 2.6%.

¹³Since there is no labor-leisure choice, we need not differentiate between a lumpsum and a proportional income tax.

Conditional on using parameters that satisfy the criteria for multiplicity, we find that modest rates of investment subsidy, 4 to 7% for most cases, are adequate to trigger development in an economy stuck in the stagnant equilibrium. Has there been any sSA economy that has successfully developed by following policies of market expansion, simultaneous industrialization, and investment tax credit or subsidy? The economy of Mauritius was languishing until 1970, following policies of import substitution. The establishment of export processing zones (EPZs) in 1970, with tax incentives, exemptions from import duties, and preferential credit facilities, boosted the economy, increased investment, and provided global markets to Mauritian firms, especially in textiles. The average annual growth rate between 1971 and 1977 was 8.3%. The Mauritian economy rebounded from a slowdown during 1978-1983, to record annual real output growth of 7% during 1984-1988, and growth rates of close to 6% during the recent years. In 1991, manufacturing was 23.3% of GDP, with EPZs alone accounting for 12.1%. Exports of manufactured goods rose from a negligible share of all exports in 1961 to 67% in 1991, nearly all of it from EPZs. Mauritius' tax code has been characterized by generous investment tax credits for industrial, manufacturing, shipping, and tourist activities, permitting, for instance, a deduction from income tax equal to 30% of the cash paid up as share capital.¹⁴ By 1998, Mauritius had grown enough to have a per capita GDP of \$8,236, more than ten times the per capita GDP of the worst-performing sSA countries. Even though the MSV models consider closed economies, the Mauritian drive toward expanding markets and increasing economies of scale by promoting exports, especially via investment incentives, are in the spirit of the Big Push policies.

2.5 Discussion

We find that empirically reasonable parameters yield multiple equilibria in the MSV model. Even if only one equilibrium obtains, provided it is stagnation, the policies discussed in the previous subsection would be relevant. In the context of static and two-period models it is a bit tricky to address the issue of whether a one-time policy intervention or a continuous one is needed to get an economy out of stagnation. The expectational nature of multiplicity and the types of policies needed to break the “bad” expectations leads us to interpret the policy intervention as one-shot. This would be especially true if technological regressions after an economy becomes industrialized are rare in a multi-period version of the model, and are slow when they do occur.

A narrow view of the fixed costs in MSV would identify them only with technological costs; however, a broader view would include the costs of regulation. Regulation costs of starting a business, which are 224.2% of per capita income in sSA, but only 8.1% for OECD

¹⁴See Lamusse (1995), *Mauritius* at <http://www.taxhavenco.com/osm/taxhavens/Mauritius.html>, and *Structural Transformation of the Mauritian Economy: 1960s - Beyond 2000*, at <http://ncb.intnet.mu/medrc/beyond.htm> for details.

countries and 11.3% in the prosperous Botswana even within sSA, could also be potentially identified with these fixed costs.¹⁵ The huge costs seen for sSA imply that only the stagnation condition would be satisfied for all three MSV models. This suggests a complementary policy intervention, namely regulatory reforms to ease entry costs. However, estimating the cost of such reforms would be considerably harder than the tax and subsidy analysis we conduct.

3 Occupational Choice

We consider the work of Banerjee and Newman (1995), in which the presence of imperfect capital markets and heterogeneity in wealth affect occupational decisions of agents, and hence economic and institutional development. We focus on a particular example, in which both a stagnant and prosperous steady state are possible; if the initial distribution is tilted toward the poor, with a small measure of middle-income agents, stagnation can result.

3.1 Model

Banerjee and Newman (BN) consider a two-period overlapping generations setup with a continuum of agents of measure one.¹⁶ Agents derive utility according to the function $c^\gamma b^{1-\gamma} - z$, where c is consumption, b is bequest given to the child, and z is labor expended. If income is y , the indirect utility is $\delta y - z$, where $\delta \equiv \gamma^\gamma (1 - \gamma)^{1-\gamma}$. There are four possible occupations: (1) Subsisters who derive return from a “backyard” technology, which has gross return $\hat{r} < 1/(1 - \gamma)$. (2) Workers, who are hired by entrepreneurs at the competitively determined wage v (subsisters are viewed as potential workers whose services are not in demand). (3) Self-employed agents, who require I units of capital to start a project with random gross return $r : r_0$ with probability $(1 - q)$ and r_1 with probability q , with the mean return denoted by \bar{r} . (4) Entrepreneurs, who can manage $\mu > 1$ workers, *each* needing I units of capital. The random gross return is r' with the same mean return $\bar{r} : r'_0$ with probability $(1 - q')$ and r'_1 with probability q' . The worker / subsister group is denoted by L, self-employed by M, and entrepreneurs by U – the lower, middle, and upper income groups respectively. An individual’s state is w , the bequest given by the parent, while the aggregate state is $G_t(w)$, the distribution of wealth.

Self-employed agents and entrepreneurs need to borrow to finance their projects. Enforcement is imperfect. Any agent who puts down a collateral of w and borrows L , can run away forfeiting collateral, but will get caught with probability π , and suffer a monetary punishment of F . Therefore, loans made satisfy $L \leq w + (\pi F/\hat{r})$.

The measures of the agents in the three income groups are denoted by p_i , $i \in \{L, M, U\}$. Entrepreneurs demand a total amount of labor of μp_U , while the maximum supply of labor

¹⁵See Djankov, La Porta, Lopez-de-Silanes, and Shleifer (2002).

¹⁶BN introduce the notion of economically active agents, but set this measure to one eventually.

by workers is p_L . Only two equilibrium wages are possible. The low wage of $\underline{v} = 1/\delta$ is the minimum wage needed to induce subsisters to work and results when $p_L > \mu p_U$. The high wage of $\bar{v} \equiv ((\mu - 1)/\mu) I(\bar{r} - \hat{r})$ is the maximum wage that will leave the entrepreneurs indifferent to being self-employed instead, and results when $p_L \leq \mu p_U$.

Given the capital market imperfection, occupational choice is driven by wealth thresholds. Agents with wealth $w \in [0, w^*]$, where $w^* = I - (\pi F/\hat{r})$, are workers (but if wage is \underline{v} , the labor market clears by some workers idling). Those with $w \in [w^*, w^{**}]$, qualify for a loan to finance self employment, where $w^{**} = \mu I - (\pi F/\hat{r})$. Finally, agents with $w \in [w^{**}, \bar{w}]$, where \bar{w} is the highest possible wealth level that can be sustained in the long run, qualify to become entrepreneurs (but if wage is \bar{v} , they are indifferent to being self-employed, and the labor market clears by p_L/μ becoming entrepreneurs and the remaining $p_U - p_L/\mu$ staying self-employed). It follows that $p_L = G_t(w^*)$, $p_U = 1 - G_t(w^{**})$, and $p_M = 1 - p_L - p_U$.

The bequest given from current income induces the distributional dynamics in wealth. That is, $w_{t+1}(w_t) = b_t = (1 - \gamma) y_t(w_t)$. This is not a linear system since the transition *rule* itself changes depending on the current distribution and therefore the equilibrium wage. However, this wage takes only one of two values, \underline{v} and \bar{v} . Moreover, attention is restricted to parameter configurations that yield tractable transition functions. If *every* starting wealth level within a given income group for a given realization of the return implies a transition into a *single* income group in the next period – for example, children of *all* the M -agents who have a good realization this period start next period as U -agents – then the two state variables, p_L , p_U , are sufficient statistics for the distribution.

We focus on BN’s example of prosperity and stagnation.¹⁷ This case results when self-employment earnings have a large spread and entrepreneurial spreads are even larger. When the low wage prevails, the low income state is absorbing; bad realizations in the middle and upper income states can push their next generations into this absorbing state. When the wage is high, the low income state allows escape into the middle income group and through it to the upper income group for good return realizations. Therefore, movements from the middle and upper income groups to the lower income group caused by bad return realizations are purely transitory. The initial presence of a substantial measure of middle-class agents is crucial for a good long-run outcome. The higher measure of middle-class agents implies a lower measure of poor agents, increasing the chance of a high wage economy with the concomitant benefits of transition described above. Moreover, the high mobility of the middle-class can increase the measure of entrepreneurs and the wage over time even when starting from a low-wage situation. Given this, the authors trace out the dynamics for two scenarios. If the starting ratio of poor (L) to wealthy (U) is low, or high but with a lot of middle-class agents (M), the prosperous steady state will be reached. However, if the L to U ratio starts off high, with few M -agents to begin with, both the U and L -agents grow at

¹⁷The transition function for this example is given in their Figure 4 and the phase diagram in Figure 5. The differential equations for p_L and p_U are given in equations (6) and (7).

the expense of the M -agents and the economy collapses to stagnation. Which steady state the economy ends at depends exclusively on the initial wealth distribution.

3.2 Calibration

As with the MSV models, here too we find parameters that work and then assess them for empirical plausibility.¹⁸ To capture the spirit of the BN model, we assume that all the model parameters are structural (invariant) across countries and differences in long-run attainment result only from differences in the initial distribution of wealth.

We assume each model period (generation) is 20 years.¹⁹ The gross subsistence or risk-free return is, $\hat{r} = 1.1$, which translates into an annual return of 0.48%. According to Dimson, Marsh, and Staunton (2002), average annualized real return on long-term bonds across sixteen countries during 1900 through 2001, was 0.7%. So the assumed figure appears reasonable. The utility parameter γ is set to 0.9, which results in $\delta = 0.72$. This value of γ implies an intergenerational persistence in the model of $(1 - \gamma)\hat{r} = 0.11$. While early estimates of this parameter in data were in the 0.2 to 0.25 range, according to Stokey (1998) later estimates, which correct for problems in the data, are in the 0.5 to 0.6 range. These are much larger than the value implied by our parameters. The “span of control” parameter μ is 2.2. Ortin-Ángel and Salas-Fumás (2002, Table 2) estimate the log of span of control to be between 1.024 and 1.5642 for a general manager, depending on the functional area – the value we use is in the ballpark of the 2.78 figure implied by the lower end of the above range.

We use $r'_0 = 1.3, r'_1 = 10.2$, for entrepreneurial returns, which translate to annual bad and good returns of 1.32% and 12.31%. These appear plausible given stock market returns.²⁰ We use a probability of the good outcome of $q' = 0.4607$ which yields an average return of $\bar{r} = 5.4$. Annualized, this is 8.8%, which is a bit higher than the 6.3% historical return presented in Burtless (1999). The self-employed returns we need are $r_0 = 1, r_1 = 18.6$, which in annual terms are 0% and 15.74%. The probability of the good outcome q has to be set to 0.25, to equate the mean returns for both types of project. If one interprets entrepreneurial (large project) and self-employed (small project) returns as the returns to public and private equity respectively, the higher spread for self-employed returns is consistent with the higher

¹⁸The strategy of starting with data-driven parameters does not allow us to replicate the prosperity versus stagnation example in a way that preserves the tractability of the BN model. However, it is important to note this does not rule out multiple steady states in a more general setup; we do not pursue the computation of such a setup as it would lead us far afield of BN’s treatment.

¹⁹This choice mainly plays a role in interpreting the project returns in annual terms and transition times in years.

²⁰See, for instance, Chart 1, in Burtless (1999), which conveniently presents 15-year average annual returns from 1871-1998. While the high value we use corresponds quite closely to his data, the low value is higher than the slightly negative return he obtains. Note that the model constrains all returns to be positive.

dispersion for private equity reported by Moskowitz and Vissing-Jorgensen (2002).²¹

We interpret \underline{v} and \bar{v} as the wages in poor and rich countries, anticipating their steady states. The ratio of wages, $p \equiv \bar{v}/\underline{v}$, is set to 5. Data on nominal wage differences across individual countries are too widely dispersed to be of use in our calibration. Ashenfelter and Juradja (2001, Table 1) compute real wages in terms of Big Macs per hour of work across 27 countries at vastly different levels of development. Calculating the averages of these wages in the top and bottom quartiles, we obtain a PPP-adjusted wage differential of 7.4, which is in the ballpark of the p used. Using the expressions for the wages, we can back out $I = p\mu / (\delta(\bar{r} - \hat{r})(\mu - 1))$, which yields, $I = 2.95$.

Let x denote the minimum fraction of a loan needed as collateral. Since w^* is the minimum wealth needed to qualify for a loan, we write $w^* = xI = I - (\pi F/\hat{r})$, which in turn implies $\pi F = (1 - x)I\hat{r}$. We require $x = 0.225$, which yields $\pi F = 2.515$; we do not need to pin down π and F separately. Using this in the expressions for the thresholds, we get, $w^* = 0.6638$, and $w^{**} = 4.2054$. We compute the maximum possible wealth, \bar{w} , as 6.3. The collateral to loan ratio at w^{**} (for the marginal entrepreneur) is $w^{**}/\mu I = 64.8\%$. As a point of empirical contact, The *Fed's Survey of Terms of Business Lending, 2004*, reports that the percentage of value of commercial and industrial loans made by domestic banks, which we interpret as entrepreneurial loans, secured by collateral is 65%.

In summary, while the parameters that are able to replicate the stagnation and prosperity example need to be rationalized ex post, most of them appear empirically relevant, with the preference parameter the least plausible. The parameters used are presented in Table 2.²² A low steady state of $p_L^* = 1$, $p_M^* = p_U^* = 0$ (stagnation) and a high steady state of $p_L^* = 0.4063$, $p_M^* = 0.4063$, $p_U^* = 0.1873$ (prosperity) result.²³

²¹Their use of “entrepreneurship” differs from BN’s use of the word. BN connect “factories” with entrepreneurs and “cottages” with the self-employed. Therefore, it appears reasonable to connect the private equity of Moskowitz and Vissing-Jorgensen (MVJ) to the self-employed returns of BN. MVJ note that the “average return to private equity is similar to that of public equity,” which is consistent with the BN assumption of equal average returns for both types of projects.

Incidentally, compounded real returns we compute from MVJ’s minimum and maximum nominal returns for a cross-sectional distribution on public equity are in the ballpark of the r'_0 , r'_1 we use. While their lowest private equity return is negative in the cross-section, we assume a value of zero for r_0 , the lowest return consistent with model assumptions. The value we use for r_1 is close to their 3rd quartile return.

²²Even small deviations from this combination of parameters cause the tractability of the transition functions that permit the characterization of prosperity versus stagnation to break down. It should also be noted that we fail to match the BN transition function for the high wage in one respect that does not seem crucial. With the bad realization, entrepreneurial incomes are negative, even though expected incomes are positive. We, therefore, need to assume an insurance scheme, presumably funded by the government from lumpsum taxes, that will cover losses and leave the children in the low-income category next period with zero rather than negative wealth.

²³The measure of entrepreneurs in the high steady state (18.73%) is higher than the 12% number reported, for instance, by De Nardi and Cagetti (2003), but in the same order of magnitude.

Table 2: Parameters for the Banerjee-Newman Model

Parameter	Value	Comment
γ	0.9	Utility parameter; implied intergenerational persistence of 0.11; compare with Stokey (1998)
\hat{r}	1.1	Annual risk free rate of 0.48% compounded over 20 years; compare with Dimson et al. (2002)
μ	2.2	Span of control; Ortin-Ángel and Salas-Fumás (2002)
r'_0	1.3	Entrepreneur's low annual return of 1.32% compounded; compare with Burtless (1999)
r'_1	10.2	Entrepreneur's high annual return of 12.31% compounded; compare with Burtless (1999)
q'	0.4607	Probability of high entrepreneurial return; determines average return \bar{r} below
r_0	1	Self-employed low annual return of 0%; compare with Moskowitz and Vissing-Jorgensen (2002)
r_1	18.6	Self-employed high return of 15.74%; compare with Moskowitz and Vissing-Jorgensen (2002)
q	0.25	Probability of high self-employed return; to get average return \bar{r} below
\bar{r}	5.4	Annual average return of 8.8%; compare with Burtless (1999)
I	2.95	Backed out from rich-poor wage ratio of 5; Ashenfelter and Juradja (2001)
π^F	2.515	Implied collateral to loan ratio of 64.8%; <i>Fed's Survey of Terms of Business Lending 2004</i>

3.3 What the Calibrated Model Explains

As mentioned earlier, the BN model explains prosperity versus stagnation based on initial income distribution. Can we find examples of sSA countries that can illustrate this? We consider the examples of Tanzania and Mauritius. Considering two economies in the same sSA region would be in the spirit of BN of holding all parameters except the initial wealth distribution constant across countries, while abstracting from institutional and other factors. The relative prosperity of Mauritius was discussed in Section 2. In contrast, Tanzania had a PPP adjusted per capita GNP of only \$483 in 1998. Can differing distributions of income in the two countries during prior years explain how a part of this difference could have arisen? Answering this requires two sets of details to be addressed – the computation of the dynamic paths, and the mapping of empirical initial distributions to the model measures p_L , p_M , p_U .

There is a system of two linear differential equations in p_L and p_U for each of the two wage regimes. Exact solutions can be computed to these linear systems.²⁴ In practice, we start from the neighborhood of a given steady state and work backward by reversing the original differential equations. By setting the initial deviations from the steady state appropriately, one could, in principle, trace out *all* the paths that lead to this steady state.²⁵ Computation

²⁴The exact solution is of the form:

$$\begin{bmatrix} \hat{p}_L \\ \hat{p}_U \end{bmatrix} = a_{1i} V_{1i} \exp(\lambda_{1i} t) + a_{2i} V_{2i} \exp(\lambda_{2i} t), \quad i = L, H.$$

where V_i is the eigenvector corresponding to the eigenvalue λ_i . Each system has two negative eigenvalues. The constants a_{ji} are pinned down by the initial conditions for p_L and p_U . The hat notation refers to deviations from the steady state values. The MATLAB program used for the computations are available from the authors on request.

²⁵If we instead work forward from a given initial condition, we will not know the steady state we will end

of several transition paths confirms the dynamic behavior summarized earlier. A substantial measure of middle-income (self-employed) agents is needed to set the economy on the path toward prosperity.

While computing transition paths and mapping a *given* initial condition to a steady state can be done entirely in terms of the summary distribution statistics, p_L , p_U , using income distribution data to first back out the initial conditions of an economy requires a knowledge of the entire wealth distribution, which the model does not track. Therefore, we make the simplifying assumption that the entire mass of agents in a given wealth interval is concentrated at the midpoint of the interval: p_L at $w^*/2$, p_M at $(w^* + w^{**})/2$, and p_U at $(w^{**} + \bar{w})/2$. We use poverty headcount from the *World Development Indicators* to pin down p_L . We then solve for a p_U such that the income Gini coefficient calculated from the piecewise linear Lorenz curve of the model matches the income Gini reported in Deininger and Squire (1996).

Poverty headcounts are available sporadically and only for recent years. For this reason, we are forced to assume that the poverty measure presented for later dates is also relevant for earlier dates. For instance, for Tanzania, we assume that the percentage of people living below the international poverty line of \$2 per day of 59.7% in 1993 is also relevant for 1977, when its Gini coefficient was 0.52. Likewise, the only poverty headcount data we have for Mauritius is 10.6% in the 90s, which we assume is relevant for 1980 (a year close to the 1977 used for Tanzania), when its Gini coefficient was 0.457. Our method of mapping distribution data to model measures yields the following “initial” conditions: $p_L = 0.597$, $p_M = 0.1071$, and $p_U = 0.2959$, for Tanzania, and $p_L = 0.106$, $p_M = 0.4323$, and $p_U = 0.4617$, for Mauritius.

As one might suspect, given the high initial measure of middle-income agents, Mauritius is more likely to reach the high steady state. Indeed, we are able to compute paths from initial conditions that are very close to the above starting distributions implied by the data ($p_L = 0.5999$, $p_M = 0.1057$, and $p_U = 0.2944$, for Tanzania, and $p_L = 0.1026$, $p_M = 0.4332$, and $p_U = 0.4642$, for Mauritius), such that Tanzania heads toward the stagnant steady state and Mauritius heads toward the prosperous steady state.²⁶ Most of the convergence occurs in two generations.

up at, and the computation of deviations would have to be done by trial and error. However, hitting an exact initial condition would involve trial and error when we work backward, a process which we did not find too onerous. The computation also checks for switches in the wage regime along a path.

²⁶The ratios of GDP per capita in the prosperous to the stagnant steady state is 38, more than twice the ratio of 17 seen in data between Mauritius and Tanzania. This discrepancy could arise from our assumption that the entire mass within a wealth interval is concentrated at the midpoint.

3.4 Policy Experiments

BN note that given the multiplicity of steady states, a one-time intervention is all that is needed to get an economy to the right distribution that would take it to prosperity instead of stagnation. It is easier to consider a redistribution of start of period wealth than of end of period income. We can view this as an unexpected imposition of an estate tax once bequest decisions of the previous generation have been made. Feasibility requires that the new level of aggregate wealth does not exceed the old level. Continuing with the assumption that wealth within each level is concentrated at the midpoint, we can show that redistribution is constrained by

$$(P_{U,o} - P_{U,n})(\bar{w} - w^*) \geq (P_{L,o} - P_{L,n})w^{**},$$

where the subscripts o and n refer to the old and new distributions.²⁷ The aim is to decrease the measure of the L and U agents and increase the measure of M agents. Beyond this we can be agnostic about the exact flow of wealth across agents. The above expression evaluated at equality indicates the maximum amount of redistribution possible. When it is a strict inequality, the right hand side can be used to evaluate the actual amount of redistribution that occurs.

Consider the initial Tanzanian distribution discussed above. The smallest perturbation we could find that would get the economy on to a path of prosperity is $p_L = 0.5962$, $p_M = 0.1378$, and $p_U = 0.2693$. This increase in the measure of M -agents by 3.2 percentage points involves a redistribution of 3.9% of the initial wealth. What is the maximum redistribution possible on to a path to prosperity? It is $p_L = 0.4688$, $p_M = 0.3349$, and $p_U = 0.1964$, amounting to an increase of M -agents by 22.9 percentage points. This involves a redistribution of 27.9% of initial wealth.²⁸ This much larger redistribution would shave the transition time to the high steady state by more than a generation.

3.5 Discussion

We had to make a few simplifying assumptions to calibrate the BN model. While the parameter values used do not match exactly with their empirical counterparts, they are in the same ballpark. The multiple steady state example obtains for a constrained set of parameters. Despite these limitations, and the delicate dynamics and initial conditions, it is remarkable that the calibrated model has the potential to explain how two economies identical in *all* respects except their initial distributions, could have ended up at very different steady

²⁷This expression is derived by stipulating that the aggregate wealth, $p_L (w^*/2) + (1 - p_L - p_U)(w^* + w^{**})/2 + p_U (w^{**} + \bar{w})/2$, at the new distribution does not exceed that of the old.

²⁸This experiment involves computing multiple transition paths and choosing one close enough to the original initial condition that leads to the high steady state. We cannot rule out the high degree of sensitivity to initial conditions as the reason for the small magnitude of the minimum required redistribution. The maximum redistribution point presented is further along the path, closer to the steady state.

states. The policy conclusion that a one time redistribution of wealth can alter the path of development finds empirical support in the land reforms of China in the early 80s, which some associate with the subsequent Chinese economic development.²⁹ The experiments also indicate a trade-off between the amount of redistribution – which is, in turn, connected to political feasibility – and transition times.

4 Human Capital Accumulation

We calibrate a simple heterogeneous-agent, two-period overlapping generations model of education acquisition that exhibits stagnation when the cost of education is high relative to income.³⁰ The model is in the spirit of Galor and Zeira (1993) in that it features indivisibility of education and fixed costs. However, unlike their model, enrollment does not automatically imply success in our model; it is probabilistic and depends on ability. Besides being empirically relevant, this leads to the implication that redistribution even among the poor is capable of prying the economy out of stagnation. The Galor and Zeira (1993) setup allows one to think of redistribution from the rich to the poor, but this channel is inoperative at a stagnant steady state.³¹ Caucutt and Kumar (2004) provides a detailed theoretical and quantitative exposition of different education policies. We provide a synopsis here and focus only on a tax-and-subsidy policy and foreign aid.

4.1 Model

The economy is populated by a continuum of two-period lived agents, each generation of measure one, in an overlapping generations setup. Children differ in their ability to become educated; more generally, this captures the “functionality” of a family. Conditional on being enrolled, a child with ability a completes education with probability $\pi(a)$; with probability $(1 - \pi(a))$, the child drops out and becomes an uneducated worker. The probability function satisfies: $\pi(0) = 0$, $0 < \pi(a) \leq 1$, $\forall a \in (0, 1]$, $\pi'(a) > 0$, $\forall a \in [0, 1]$. The function π can be used to capture the quality of the educational system. The distribution function for ability on the support $[0, 1]$ is denoted by $F(\cdot)$, and $f(\cdot)$ is the corresponding density function. Ability draws are iid.

²⁹See the volume *Land Reform: Land Settlement and Cooperatives* (Special edition), 2003, published by the FAO and the World Bank for experiences and perspectives of land reform and its effect on growth and poverty reduction in several countries.

³⁰The model here extends the one in Caucutt and Kumar (2003) by modeling indirect costs explicitly, and focusing on conditions that give rise to stagnation.

³¹An external shock can shift the transition function in their paper and also in the representative agent setup of Becker, Murphy, and Tamura (1990), who additionally consider fertility decisions. We instead devise policy measures that would shift the transition function upward.

Enrolling a child involves a real cost of e_d units of consumption. A parent cannot borrow to finance her child's education. If a child is not enrolled she can work and add w_c to the family's consumption, by performing tasks such as tending livestock, fetching water, and helping in the fields, in addition to supplying labor outside the family. If the child is enrolled in school, she can contribute only φw_c to the family, where $0 < \varphi < 1$. The total cost is $e = e_d + (1 - \varphi) w_c$, the sum of direct and indirect costs. If education costs are subsidized to the level s , it is netted out of the cost e .

The aggregate state variable in this economy is the measure of educated workers entering the labor force at any period, n_e . The wage earnings of an educated parent as a function of the aggregate state is denoted by $W_e(n_e)$, and that of an uneducated parent by $W_u(n_e)$. The earnings of a household that does not enroll its child is $w_j(n_e) \equiv W_j(n_e) + w_c$, ($j = e, u$), and one that does is $w_j(n_e) - e$. Workers inelastically supply their unitary time endowment.

There is a single consumption good produced using educated and uneducated labor as inputs. The CES production function is $Y = A[\theta(N_e + \gamma N_u)^\nu + (1 - \theta)(N_u + \varepsilon N_e)^\nu]^{\frac{1}{\nu}}$, where $0 < \gamma, \varepsilon, \nu < 1$, and $\gamma < \varepsilon$.³² The first term within the square brackets can be thought of as "brain" and the second term as "brawn". Here, N_e is the number of educated workers, and N_u the uneducated workers, employed by the firm. Educated workers are the primary suppliers of "brain". The weight of uneducated workers in this factor, γ , is small and keeps wages bounded even when the economy is stagnant. The mere hiring of a particular type of worker contributes to both factors in the proportion shown above. In a competitive labor market, the wage rates W_e and W_u would be the appropriate marginal products and decreasing in N_e and N_u respectively.

A parent of a given type, who has a child of ability a , optimally decides between enrolling and not enrolling the child, weighing the utility cost of education against the possibility the child becomes educated and gets higher utility. The intergenerational discount (altruism) factor is β . The future aggregate state is posited to follow $n'_e = \Phi(n_e)$, which will be consistent with the outcome of a perfect foresight equilibrium.

This model delivers a threshold ability for enrollment; parent type j enrolls her child if $a \geq a_j^*(n_e)$, and does not otherwise. If $a_j^*(n_e) = 1$, even the most able child will not be enrolled, and the enrollment rate of type j children, given by $\left(1 - F\left(a_j^*(n_e)\right)\right)$, is zero. Enrollment rates are higher among the rich; $a_e^*(n_e) < a_u^*(n_e)$, for $n_e \in [0, 1]$.

A steady state satisfies $\Phi(n_e^*) = n_e^*$, with stagnation defined as a locally stable steady state at $n_e^* = 0$. Caucutt and Kumar (2004) show that a necessary condition for stagnation is $a_u^*(0) = 1$ (the poor do not enroll their children). This condition is sufficient if it additionally holds in a neighborhood of $n_e = 0$. When the initial fraction of educated people in the workforce is too low, the wages of the uneducated workers are too low for them to find it profitable to send their children to school. This results in a decrease in the fraction of

³²Stokey (1996) considers a similar production function.

educated workers next period, which further decreases the wages of the *un*educated workers and reinforces the above-mentioned behavior. The dynamic behavior of the economy around the origin is therefore mainly governed by the utility cost of poor parents. Provided e is not prohibitively high, rich parents always enroll a positive fraction of their children, especially near the origin when their wages are very high. But there is vanishingly small measure of them.

What conditions yield $a_u^* = 1$ in the neighborhood of the origin and hence stagnation? For a general utility function, a sufficient condition is

$$(u(w_u(0)) - u(w_u(0) - e)) > \left(\beta / \left(1 - \beta \int_0^1 \pi(a) dF(a) \right) \right) [u(w_e(0) - e) - u(w_u(0) - e)].$$

The left hand side is the *utility* cost of education to the *poor* parent. When this is greater than the discounted, maximum possible utility gain from education, stagnation results. The discount factor reflects differential enrollment rates. With an isoelastic utility function $u(c) = c^{1-\sigma} / (1-\sigma)$, $\sigma \geq 0$, it can be shown that the condition for stagnation is more likely to be satisfied when the curvature of the utility function and cost of education are high, and the wage gap and the discount factor are low.

Similar to the models considered in the previous two sections, ours is capable of delivering an additional stable steady state with a positive fraction of educated agents. This happens if the rate of enrollment of the uneducated increases rapidly when n_e increases (w_u increases). Whether an economy ends up at the low or high steady state would then depend on the starting level of n_e .

4.2 Calibration

We assume that all children are born with two years of education (“uneducated”) and successful education involves completion of a further eight years of schooling (“educated” agents thus have ten years of education). For eighteen of the worst-performing sSA countries, we use the data in Barro and Lee (1996) to find that median years of primary attainment is about 1.5 years; this motivates our baseline level of education.³³ Secondary schooling indicators are often used in cross-country growth studies and completion of education at this level is considered the minimum level needed for a worker to perform well in the modern economy; this motivates our definition of educated workers. The median years of secondary attainment in the above-mentioned sample of sSA countries is 0.15, which conforms to our view of stagnation. We assume agents are born at age 6 and are young until the age of 25; they become adults at the age of 26, have a child, and die at the age of 45. The model

³³The countries we consider are Angola, Benin, Burkina Faso, Burundi, Central African Republic, Chad, Djibouti, Guinea, Guinea-Bissau, Malawi, Mali, Mozambique, Niger, Rwanda, Somalia, Tanzania, and Uganda.

period is thus 20 years. The life-span corresponds closely to the median life expectancy of 45.5 years in this sample.

Table 3: Predetermined Parameters for the Caucutt-Kumar Model

Parameter	Value	Comment
β	0.6676	Annual discount factor of 0.98 compounded over 20 years
ν	0.35	Educated-uneducated elasticity of substitution of 1.54; Autor et al. (1998)
ε	0.1	Skilled labor counts 10% of unskilled labor towards “brawn”
σ	3.5	Utility RRA parameter; refer Ogaki and Zhang (2001)
$F(a)$	a	Uniform ability distribution

We start by assuming values for certain parameters that are commonly used in the literature. The generational discount factor is set at $\beta = 0.6676$, which corresponds to a yearly discount factor of 0.98 compounded over 20 years. We set $\nu = 0.35$, which corresponds to an elasticity of substitution between educated and uneducated labor of 1.54. Autor, Katz, and Krueger (1998) report that the emerging consensus on the elasticity between skilled and unskilled labor is approximately 1.4 to 1.5. In the absence of direct evidence, we set $\varepsilon = 0.1$, (each unit of skilled labor counts 10% of unskilled labor toward brawn) and leave $\gamma < \varepsilon$ as a free parameter. We appeal to arguments for a negative relationship between relative risk aversion and wealth (see, for instance, Ogaki and Zhang (2001)), and set the curvature parameters of the utility function, σ , at a higher value of 3.5 instead of the 2 usually assumed in calibrated macroeconomic models. We assume a uniform ability distribution in $[0, 1]$; that is, $F(a) = a$. The predetermined (“structural”) parameters we use are summarized in Table 3.

We allow for the possibility that the human capital production functions can differ across the two types of families, to account for the advantages educated families might have in the production of human capital. The parametric form we use is $\pi_i(a) = k_i (4a^3)$, in the interval $[0, 1/2]$ and $k_i (1 - 4(1 - a)^3)$ in $[1/2, 1]$. This convex-concave parametric form was chosen because it allows us to better match the enrollment and dropout rates in the vicinity of stagnation. Such a shape is not *required* to obtain a stagnant steady state. We set $k_e = 1$ and allow only k_u to vary. Therefore, the production parameters, A , θ , and γ , education parameters e_d , s , w_c , φ , and the human capital production parameter k_u need to be chosen. Since high costs for low quality education and a low wage gap are responsible for stagnation in our model, we choose these remaining parameters to specifically match the following sSA targets. We again use data from wherever it is available, and specify ranges where necessary.

1. Education attainment: We target an attainment, n_e^* , of zero. Calibrating the model to a steady state attainment close to zero instead of exactly zero will not alter the policy conclusions significantly.
2. Skill premium: Bils and Klenow (2000) present Mincer regression coefficients on schooling for a few sSA countries: 0.207 for Cote d’Ivoire, 0.126 for Botswana, and 0.067 for Tanzania.

Given the use of log wages and 8 years of schooling, we compute the corresponding premia to be 5.24, 2.74, and 1.71 respectively. When the figures reported in Bigsten et. al. (2000) are used to compute the premium for our education definition, we obtain a value of 1.42.³⁴ The premium therefore spans the rather wide range of 1.42 to 5.24.

3. Enrollment rates: A “naive” measure of enrollment rate, obtained by taking a simple average of the primary and secondary enrollment rates and averaging it across our subsample of countries, is 31.8%. In the model, education begins at the third year and continues for eight years. Using year-to-year survival rates from the *World Education Indicators*, we calculate enrollment rates conditional on surviving the first two years of education; the average of this enrollment rate is 22.9%.

4. Dropout rates: The “naive” dropout rate is 32.3% for the sample of countries, and the dropout rate conditional on surviving the first two years of education is 13.5%. This dropout rate is likely to be underestimated, since data is missing for several of the poorest countries.

5. Education subsidy as a fraction of parental cost: Ablo and Reinikka (1998), in Table 5, present data on parental and government spending in Uganda for 1991 through 1995. In conjunction with the per capita GNP figures, we compute the annual share of income that is spent on education and the parental share of this cost, averaged over 1991-95. If we denote per capita income by y , then $\lambda_1 \equiv \text{total direct cost}/y = 8.1\%$, $\lambda_2 \equiv \text{govt. cost}/y = 2.7\%$, and therefore $\lambda_1 - \lambda_2 = \text{parent's cost}/y = 5.4\%$. This implies the ratio of subsidy to direct cost of education is $s/e_d = 1/3$.

6. Indirect cost: Bredie and Beeharry (1998), in Annex A, present time use data of school-aged children in Madagascar and conclude that the opportunity cost for boys in school is 20 hours per week. This figure is in line with the 21 hours per week reported by Beegle, Dehejia, and Gatti (2002) for Tanzania. We assume this is half the adult work week; non-schoolgoing children work half an adult week and schoolgoing children work none. We impute the average wage in the economy to this time; we set $w_c = 0.5y$, where y is the average wage earnings (w_u in a stagnant economy).

7. Education expenditure to GDP: Consider family income when the child is not enrolled. The present value of the parent’s annual income y over 20 years at an 8% annual rate of discounting is $10.6y$. The present value of the child’s income is half this at $5.3y$. If the family enrolls the child, the present value of the annual parental cost of education $(\lambda_1 - \lambda_2)y$ over the eight schooling years is $6.2(\lambda_1 - \lambda_2)y$. If the child goes to school, it is assumed that after the first 8 years, the child can work the rest of his youth years with annual earnings of $0.5y$; the increased earnings due to education are not realized until adulthood.³⁵ The present value

³⁴We use their coefficients from regression (3) in Table 7 to compute wages for 2 and 8 years of education.

³⁵With this assumption, we attempt to account for the experience premium, which we have not explicitly modeled.

of these earnings is $2.2y$. Therefore, φ , the ratio of the earnings of the schoolgoing child to the non-schoolgoing child is $2.2/5.3 = 0.415$. We also calculate the direct education expenditure net of government subsidies as a fraction of GDP as $(e_d - s)/Y = 6.2(\lambda_1 - \lambda_2)y/10.6y = 0.0316$.

Table 4: sSA-specific Parameters for the Caucutt-Kumar Model

Parameter	Value	Comment
A	2	TFP; consistent with stagnation ($n_e^* = 0$)
θ	0.48	Skill share in production; θ and γ pinned down by skill premium
γ	0.05	Unskilled labor counts 5% of skilled labor towards “brain”; θ and γ pinned down by skill premium
k_u	0.85	Human capital production parameter for poor; governed by enrollment and dropout rates
e_d	0.0326	Direct cost of education; from direct education costs borne by families as a fraction of GDP
s	0.0109	Subsidy; from fraction of education costs subsidized in Ablo and Reinikka (1998)
w_c	0.3439	Child’s wage; governed by indirect cost of education in Bredie and Beeharry (1998)
φ	0.415	Earnings ratio of schoolgoing to non-schoolgoing child; calculated from Ablo and Reinikka (1998)

In Table 4, we present values for the specific parameters resulting from our calibration attempt to best match the above targets. Given that we fixed $\varepsilon = 0.1$, it is reassuring that a $\gamma = 0.05 < \varepsilon$ results from our calibration; a unit of unskilled of labor contributes less to “brain” than does a unit of skilled labor to “brawn”. Likewise, given that we set $k_e = 1$, it is reassuring that $k_u = 0.85$; the human capital production function for families with educated parents dominates the one for families with uneducated parents. Finally, the resulting value for the total cost of education, $e (= e_d + (1 - \varphi)w_c)$, is 0.234; in other words, the direct cost is only 14% of the total costs.

4.3 What the Calibrated Model Explains

With the chosen parameters, stagnation results ($n_e^* = 0$). The skill premium is 4.96, which is within the range seen in data, though close to the upper end. Exactly at the point of stagnation, there is no enrollment; $a_u^* = 1$, and even though $a_e^* = 0.12 < 1$, there is a zero measure of educated people. Therefore, we examine the average dropout and enrollment rates in the “vicinity” of stagnation ($n_e = 0.00 - 0.15$), with the interpretation that these economies are headed toward stagnation if they are not already in it. The enrollment rate is in the range of 0 to 21%, which is a bit lower than the range given in the previous subsection, but in the ballpark of the enrollment rate calculated conditional on students surviving past the second year; the dropout rate is in the range of 24 to 43%, overlapping considerably with the range seen in data.

Even though multiple steady states are theoretically possible, for the chosen calibration the transition function yields stagnation as the only steady state.³⁶ An increase in β and a

³⁶Since this is our own model, we take the liberty of not searching for a set of parameters that yields

decrease in σ results in a non-zero steady state. Our assumption that a higher wage from education does not materialize until the second period has implications for the value of the indirect cost. A higher value for φ decreases the opportunity cost of education and makes stagnancy less likely. We search for the minimum value of φ that would get the economy out of stagnation; it is 0.44 as opposed to the baseline value of 0.415.

4.4 Policy Experiments

Given the single stagnant steady state that we obtain, we only study continuous policies to revive these economies, rather than one-shot policies to change the initial measure of educated. Since the behavior of the uneducated poor, who form the vast majority, leads to stagnation, it is natural to consider a policy of subsidizing their direct, and possibly indirect, costs of education.³⁷ We assume all students, rich or poor, get a subsidy of s each, and all workers are taxed at the rate τ . The government balances its budget every period. We hold the subsidy, s , constant at the level that leads to the target steady state. The tax rate τ is varied each period so as to balance the government budget. In all experiments, we hold the π functions at their benchmark specification; we do not make any adjustment for the quality of the education system. We do not have enough data on quality, especially from this region, to calibrate changes in π . We also expect the quality of educational institutions to move upward more sluggishly than enrollment.

We use Mauritius, as we did in Sections 2 and 3, as an example of a high-performing sSA economy. Barro and Lee (1996) indicate that the percentage of population who attended secondary school in Mauritius was 36.5% in 1990 and the percentage who completed secondary school was 28.1%. While the average primary expenditure per pupil for our sub-sample decreased from \$135.6 in 1960 to \$79.8 in 1990, it increased in Mauritius from \$256 in 1960 to as high as \$544 in 1980. The public education expenditure as a fraction of GDP was higher for Mauritius in 1990 at 3.6% when compared to the average of 2.8% for the rest (which is inflated by low GDP levels). In 1980, nearly 20% of government expenditures in Mauritius went to fund education. Given these facts one could plausibly argue that education policies were at least partly responsible for its solid economic performance. Moreover, the primary enrollment rate was close to 100% in 1960, a decade before the economy started growing, when its expenditure per pupil was nearly twice as high as the other sSA countries. These facts allay concerns that education might have *followed* rather than preceded development; indeed it has been argued that the export-oriented industries, mentioned in Section 2, were

multiplicity.

The framework is rich enough to accommodate developed countries such as the US, since higher education expenditures, higher TFP, and better quality education systems are associated with higher steady states. We do not pursue this route here.

³⁷Mexico's *Progres*a program, for instance, subsidizes both the direct and indirect costs.

attracted to Mauritius because of its better educated workforce.³⁸ We therefore analyze policies for the other sSA countries that would result in a steady state close to the 30% level of educational attainment seen in Mauritius. Since our calibration indicates that at the prevailing subsidy level, the net educational cost is still too high to escape stagnation, our aim is to find a subsidy level that would shift the transition function upward and cause the economy to not only emerge from stagnation, but also result in a $n_e^* = 0.3$.

A subsidy level of 0.09, which is 38.5% of total costs, is needed to take the economy to a steady state of $n_e^* = 0.3$. It is not enough for the government to subsidize only the direct costs of education; it would have to defray part of the child's contribution to the family income that is lost by sending the child to school. At the steady state, a tax rate of 3.2% needs to be levied on all workers to meet the cost of subsidies.³⁹ Since all workers are taxed at the same rate, the ratio of government expenditure to GDP will also be 3.2%. This is close to the 3.6% figure cited earlier for Mauritius and thus appears to be an achievable target. The closeness of the model outcome to data from a country that was not originally part of the calibration, lends support to the validity of the calibrated model and the use of education subsidies as a development policy. The expenditure to GDP of 3.2% can be put in perspective by considering the military expenditure as a fraction of GNP, which was 3.1% in sSA in 1992. Evidently, diverting part of these expenditures to education will go a long way toward meeting subsidy expenses before new taxes become necessary.⁴⁰

Each agent would have to be given 20.3% more consumption every period in the stagnant state, in order to equate an aggregate welfare measure – weighting generations within a period equally and using a discount factor of β for future generations – to that in the new steady state. When the costs of transition (increased taxes and educational investment when uneducated workers' wages are still low) is taken into account, the gain in welfare in terms of equivalent consumption of 6.6% is much lower, but still significant. The economy is very close to the steady state in four to five model periods. A tax rate of 7% maximizes welfare, inclusive of transition, and results in a steady state education attainment of 38%.

³⁸See, for instance, Anker et. al. (2001) and also Lamusse (1995), who attributes the success of the export processing zones to “a reserve labor force of literate women who were readily trainable for semi-skilled production jobs...”

³⁹This policy increases output by close to 50%. The ratio of the subsidy to per capita GDP is 8.75%. The skill premium drops considerably, to 1.54. Given the skill premium of 2.1 in Malaysia, a country of comparable educational and economic development, calculated from Bils and Klenow (2000), the drop in premium appears to be overstated. The economywide enrollment rate is 36.2%, which masks the relatively high enrollment of 61% for educated parents. The dropout rate is close to 20%.

⁴⁰See *2000 World Development Indicators*, Table 5.7, for military expenditure data.

4.5 Discussion

The redistribution motive in our setup arises from the indivisibility of education, liquidity constraints, and the focus on aggregate welfare. If the return to education falls a bit short of the amount required for enrollment for every student, aggregate welfare could be improved by redistributing and making the return attractive at least for the most able students. At the stagnant steady state everyone is poor, and it is redistribution across ability rather than income levels that gives the initial kick-start to the economy. It is important to note that the child’s ability need not be observed by anyone other than the parent; the in-kind nature of the subsidy would automatically attract the more able students. This implicit, rather than explicit, redistribution across abilities makes such a scheme politically feasible.

Unlike the models considered in Section 5, ours is not a model of political economy. However, anticipating the effects studied in those models, we speculate on the question of why such subsidy schemes might not be seen in practice. There is a drop in wages of the educated of about 61% going from stagnation to the Mauritius-like subsidy level. In the neighborhood of stagnancy, we find that these workers actually prefer stagnancy. Across all steady state comparisons, the currently uneducated prefer subsidies more than the currently educated.⁴¹ There is therefore an incentive for the educated “elite”, who often occupy key policy making positions in these countries, to *not* subsidize education and preserve the monopoly they enjoy for their children who are more likely to be educated. If this incentive effectively causes subsidization to be blocked, the economy will remain stagnant.

The tax-and-subsidy experiment suggests that even an economy locked into stagnation need not be dependent on foreign aid to trigger development. When we do conduct an experiment that mimics foreign aid, with foreign aid paying for subsidies instead of local taxes, welfare is obviously higher. But it appears inconceivable that countries will be willing to donate the foreign aid of 6.8% of GDP (10.4% relative to pre-subsidy levels) that is needed to take the economy to the same educational attainment as under optimal taxation, year after year. An intermediate strategy would be for the government to borrow, if possible, on a long-term basis from other countries or development agencies to finance increased education expenditures and alleviate transitional costs.

5 Political Economy

We consider two models, which follow quite different approaches, in this category. Bourguignon and Verdier (2000) assume that only an educated-wealthy minority (oligarchy) initially participates in the political system. An education externality provides the oligarchs an incentive to subsidize the education of the poor. However, this is tempered by the concern

⁴¹Even the poor who do not enroll their children will support the tax and subsidy scheme. Even though the unenrolled child will be poor next period, the wage of the uneducated increases.

that the newly educated poor may gain political control, and impose costly redistributive taxes.

In Acemoglu and Robinson (2002), a ruler who is assured of his position does not block technology adoption because he reaps a percentage of the increased output. But if the probability of getting replaced increases with adoption, he may choose to block the technological advance in order to remain in power. Rulers with a very firm grip on power, as well as those with a high probability of being replaced, choose to innovate, while those in the middle choose to block technology adoption.

5.1 The Politics of Education

5.1.1 Model

In the Bourguignon and Verdier (BV) model, an individual lives for two periods, young and adult. Young agents differ in the innate human capital they inherit directly from their parents. Initially there are two types of young agents, those with educated (rich) parents and those with uneducated (poor) parents. Their earnings are given by, $y^i + \mu(1 - p)$, $i = p, r$, where y^i is the inherited human capital, $1 - p$ is the fraction of young whose parents were educated, and μ is the intensity of the education externality.

A young person decides whether or not to get educated, given a cost of c in the current period, and an increase in next period's income of R . A young agent who chooses education consumes $y^i + \mu(1 - p) - c$, while young, and $y^i + \mu(1 - p + e) + R$, while old. Here, e , is the fraction of "newly educated" young – those with uneducated parents who choose to get educated. If the agent chooses no education, he consumes $y^i + \mu(1 - p)$ while young, and $y^i + \mu(1 - p + e)$, while old. With linear utility and no discounting across periods, the condition for choosing education is, $c < R$. BV assume that there is no borrowing, and that $R > c > y^p + \mu(1 - p)$. Therefore, the young with educated parents can afford education, and others cannot.

While young, rich children decide whether to subsidize the education costs for a measure e of poor children. Doing so would increase every adult's income by μe . All educated adults vote over a proportional tax rate that is redistributed in a lump sum fashion. Since newly educated adults have a lower income, these new members of the oligarchy may favor redistribution. The upper bound on redistributive taxes that can be levied on each rich agent is assumed to be τ^+ .

The authors' dynamic framework yields one of three outcomes; here, democratization is equivalent to the education of the poor.

(1) The economy never democratizes, and the measure of the oligarchy remains at the initial $(1 - p)$ forever. This is a trap. For this to occur the externality must be low relative to the cost of education: $\mu < (c - y^p) / (2(1 - p))$.

(2) The economy democratizes fully in two generations. This occurs for intermediate values

of the externality, μ . There are two further possibilities. (a) If the initial inequality is high enough, the oligarchy chooses to educate only $e = (1 - p)$ of the poor in the first generation, so that the oligarchy does not lose its voting advantage. In the second generation there are $2(1 - p)$ educated, and the poor now get enough of a boost from the education externality that they are no longer credit constrained. Everyone is educated in all future generations.⁴² (b) If the initial inequality is low enough, the oligarchy chooses to educate $e > 1 - p$ poor in the first generation. In the first generation, the oligarchy does lose its voting advantage; however, the chosen redistribution is zero. In the second generation, there are enough educated agents, and the poor get enough of a boost from the education externality that they are no longer credit constrained. Again, everyone is educated in all future generations.

(3) The economy democratizes fully in one generation. If the externality is high enough relative to the initial inequality, the oligarchy chooses to educate all of the poor in the first generation. Everyone is educated in all future generations.

The parameters of the model need to satisfy a few constraints. The education externality, μ , and the earnings $y^p + \mu(1 - p)$ and $y^r + \mu(1 - p)$ must be positive. The return to education must be greater than its cost, so that at least the rich children would attend school. The cost of becoming educated must be greater than the poor's earnings, $c > y^p + \mu(1 - p)$, for education subsidies to be relevant.

For analytical tractability, BV assume linear preferences; the poor children do not attend school only because the cost is strictly greater than their income. From the human capital model studied in Section 4, we find that the concavity of preferences plays a big role in education decisions. While assuming concave utility makes the model analytically intractable, we can perform numerical exercises using calibrated parameters. Therefore, we also consider a lifetime concave utility specification with discounting of, $\log(c_1) + \beta \log(c_2)$.

With concave utility, the young person equates the marginal utilities of consumption while young and old, and the subsidy required is higher than just the difference between cost and income. The subsidy is also increasing in the fraction of newly educated people, e . For a poor person deciding on schooling, an increase in e implies an increase in his adult income and a decrease in his marginal utility of consumption. A commensurate increase in consumption while young is needed to equate marginal utilities. Therefore, a larger subsidy is needed to induce enrollment. With linear utility, the education externality can boost future income of the poor enough to ease the liquidity constraint. With curvature, the externality can initially work against the subsidy, causing *fewer* agents to get educated in the first place. There is another effect that could decrease subsidization by the oligarchs. Since they are relatively richer in their adult period, they are even less willing to pay a cost while young to reap a benefit when old, an effect that is magnified by discounting.

⁴²The conditions for this and other remaining cases are algebraically involved and can be found in the paper.

5.1.2 Calibration

We remain as close to the spirit of the model as possible while calibrating this repeated-two-period model. This is not necessarily an overlapping-generations model, since generations do not interact aside from transmitting human capital “genetically”. Therefore, the two periods need not be of the same length; the period an agent is young need only be as long as it takes to acquire education. With linear preferences, we follow the authors and do not discount within or across periods, but we do so with concave preferences. This makes the mapping of yearly data, which we do at the beginning of the next subsection, into the youth and adult periods straightforward. Because the economy starts with only the rich being educated, $1 - p = e$, initially. The parameters we need to calibrate are $1 - p$, y^p , y^r , R , μ and c . Here, μ is treated as “structural”, and the income distribution and cost parameters are necessarily chosen to target an sSA country, Tanzania in this instance.⁴³

In Tanzania, 9.6% of the population completed primary school in 1990, where primary school is of a 7-year duration. Therefore, we set the fraction educated, $1 - p$ to 0.1 and s , the number of years of schooling to 7.

Given the average adult income, A_I , and the percentage of income earned by the uneducated, G , we write two conditions that simultaneously yield $y^p(\mu)$ and $y^r(\mu, R)$:

$$\begin{aligned} A_I &= (1 - p)(y^r + \mu(1 - p) + R) + p(y^p + \mu(1 - p)) \\ G &= p(y^p + \mu(1 - p))/A_I. \end{aligned}$$

From the World Bank’s *World Development Indicators* (WDI), we use the per capita Gross National Income (GNI) for A_I . For Tanzania, in 1993, $A_I = \$170$. We set $G = 0.699$, the fraction of income earned by the poorest 90% in Tanzania in 1993, as reported by WDI. The above two equations then imply $y^p(\mu) = 132.03 - 0.1\mu$, and $y^r(\mu, R) = 511.70 - R - 0.1\mu$.

We select μ using estimates of the percentage increase in *average* log earnings due to increasing *average* education by one year, denoted by D . We first write

$$D = \frac{\partial \log A_I}{\partial (1 - p)_{data}} = \frac{\partial A_I}{\partial (1 - p)} \frac{\partial (1 - p)}{\partial (1 - p)_{data}} \frac{1}{A_I}.$$

Since there is a difference between the average population getting one more year of schooling (in empirical estimates) and one person getting s years more of education (in the model), we make the distinction between $(1 - p)_{data}$ and $(1 - p)$. Noting this, and the expression for A_I , we simplify the above equality to

$$D = \frac{(y^r(\mu, R) - y^p(\mu) + R + \mu) \frac{1}{s}}{A_I}.$$

Substituting in for $y^p(\mu)$, $y^r(\mu, R)$, A_I , and s yields $\mu = 1190D - 379.67$.

⁴³The cost and income parameters were specific to the sSA economies in the human capital model as well.

Heckman and Klenow (1997) report that in log per capita GDP regressions, the coefficient on average years of schooling attained in the population age 15 and above, is 0.3. When the authors control for technology, the estimate drops to 10.6% in 1985 and 7.0% in 1960. Venniker (2000) surveys the literature to find there is no conclusive evidence in favor of, or against, education externalities. An estimate of $D = .3$ implies a negative μ , which the model does not permit. Therefore, we use $D = 0.35$, which is slightly higher than the highest empirical estimate seen.⁴⁴ This implies $\mu = \$36.83$, and therefore, $y^p = \$128.35$ and $y^r(R) = 508.02 - R$.

We compute the model R using Mincer coefficients. Given earnings of an uneducated rich child of $y^r + \mu(1 - p)$, earnings of an educated child of $y^r + \mu(1 - p) + R$, and an estimate of the gross percentage increase in earnings from getting 7 more years of education, R_{data} , we can write, $(y^r + \mu(1 - p))R_{data} = (y^r + \mu(1 - p) + R)$. Hall and Jones (1999) use a return of 13.4% from the first four years of education, 10.1% from the next four, and 6.8% from the next two. With $s = 7$, $R_{data} = 2.2071$, and $R = \$279.86$. Therefore, $y^r = \$228.16$.⁴⁵

As in Section 4, we assume that indirect costs are half of the earnings of a poor (uneducated) person. Penrose (1998) reports direct costs of education in Tanzania. The government spending per primary pupil in 1994/1995 was \$20.51. The maximum reported spending by households on a primary pupil in 1994 was \$79.90, while two-thirds reported spending between \$8.28 and \$31.05.⁴⁶ Therefore, an upper bound on total cost is $c_1 = \$20.51 + \$79.90 + .5(y^p + \mu(1 - p)) = \166.43 . Using the upper end of the range in which two-thirds of the respondents fall, $c_2 = \$20.51 + \$31.05 + .5(y^p + \mu(1 - p)) = \117.58 . We summarize the parameters we use in Table 5.

Table 5: Parameters for the Bourguignon-Verdier Model

Parameter	Value	Comment
μ	\$36.83	Externality parameter; compare with Heckman and Klenow (1997)
$1 - p$	0.1	Fraction of population completing primary school; Tanzania 1990
y^p	\$128.35	Earnings of poor; from WDI data on per capita income and poorest 90% income
y^r	\$228.16	Earnings from rich; from WDI data on per capita income and poorest 90% income
R	\$279.86	Monetary gain from education; from Mincer returns in Hall and Jones (1999)
c	\$117.58, \$166.43	Lower and upper bounds on cost of education; Penrose (1998)

⁴⁴Other estimates for education externality are even lower. Rauch (1993), without correcting for the endogeneity, finds that a one year increase in average education raises wages by 3 to 5 percent in 1980. Acemoglu and Angrist (2000) use child labor laws and compulsory schooling laws as instruments for average schooling, and find external returns to education of around 1 percent that are not statistically significant.

⁴⁵Bils and Klenow's (2000) Mincer coefficient for Tanzania is 6.7% per year. The implied $R_{data} = 1.5745$, and $R = \$186.71$, are in the ballpark of the estimates implied by Hall and Jones (1999).

⁴⁶A 1994 exchange rate of .00207 \$/Schillings from the WDI is used in these calculations.

5.1.3 What the Calibrated Model Explains

The values in Table 5 imply that the poor young and poor adults earn $y^p + \mu(1 - p) = \$132.03$. The earnings of the young rich are $y^r + \mu(1 - p) = \$231.84$, and the adult rich are $y^r + \mu(1 - p) + R = \$511.70$. The education “premium” is 3.88, which lies in the range of the estimates reported in Section 4. We map the yearly data into values for the young period and the adult period. We assume agents are young for 7 years and old for 33 years. This implies that, with linear preferences, earnings of the young and poor are \$924.01, young and rich \$1622.88, adult and poor \$4356.99, and adult and rich \$16,886.10. The increase in earnings from schooling is \$9235.38, the education externality is \$1215.39, and the cost of education is either $c_1 = \$1165.01$ or $c_2 = \$823.06$. With linear preferences, the second cost estimate is too low; even the poor young can afford education. Therefore we use the first cost estimate. The increase in earnings due to education is greater than both of our cost estimates. The condition for a trap holds since, $\mu = \$1215.39 < (c - y^p) / 2e = \1332.80 . So, without further intervention, this economy will remain with 10% of the population educated forever. A subsidy of \$241 is needed to cover the difference between costs and the income of the poor.

When preferences are concave, we discount across periods and across years within a period. We use an annual $\beta = 0.98$. This means that we need to recompute the values for earnings, costs, and returns. We find that earnings when young and poor are \$870.57, young and rich \$1528.69, adult and poor \$3212.25, and adult and rich \$12,449.52. The increase in earnings from schooling is \$6808.92, the education externality is \$896.06, and using the lower estimate of costs, the cost of education is \$701.70. With these parameters, the rich optimally choose to go to school. It turns out that the poor choose not to attend school even though they can afford to pay the lower of the two cost estimates. The result is stagnation. In order to get educated, a poor person needs a subsidy of \$155.37. This is lower than the \$241 figure for the linear case because we have assumed the more reasonable lower cost estimate.

5.1.4 Policy Experiments

When preferences are linear, the young rich are willing to pay a subsidy of \$121.54 towards educating 86.4% young poor. This is calculated from the condition for case 2b discussed earlier, where the young rich choose to educate just enough poor, so that even though they lose their voting power, the newly educated old vote for zero redistribution. However, this amount does not cover the required subsidy of \$241. If foreign aid makes up the difference of \$119.26, 86.4% of the young poor will be educated. The contribution by the rich works to 5.68% of their lifetime income. What happens to the poor who do not get educated initially? In the next generation, the education externality increases everyone’s earnings. Young people’s earnings increase by \$222.83, but there is still a shortfall of \$18.17 to cover costs of education. However, foreign aid is not required to cover this shortfall. The newly educated

members of the oligarchy are now in the majority and are not subject to redistribution. Therefore, they are willing to subsidize the remaining poor. After the second generation there will forever be three earnings levels, the initially rich, the first group of poor to be educated, and the second group of poor to be educated.

Alternatively the foreign organization could initially pay the entire subsidy for all of the poor in the first generation. If they do so, everyone will remain educated for all future periods. In this scenario there will forever be two earnings levels, the initially rich, and the initially poor. As a percentage of GDI in Tanzania, the subsidy costs for the foreign organization is 8.7% for seven years when it shares costs with the local wealthy elite, and 20.2% for seven years when it bears the entire cost. If the foreign organization can overcome problems associated with lending resources for education purposes, a lower cost solution would be to lend to the poor while they are young, and collect from them once they are old and educated.

With concave preferences, if foreign aid is low (below \$150/student), the poor continue to choose no enrollment, and the rich no subsidization. If foreign aid is \$155/student, the rich choose to send 4.5% to school and are willing to contribute \$2.76/student towards the subsidy, 0.08% of their income. For reasons explained earlier, as the fraction educated rises, the subsidy required by the poor also rises. If foreign aid is \$160/student, the rich choose to send 9.28% and contribute \$0.28/student towards the subsidy, 0.01% of their income. When foreign aid is greater than this, more students choose to get educated, but the rich do not contribute to the subsidy. When foreign aid is \$200/student, all the remaining poor choose to get educated. Clearly, these policies are much costlier for a foreign donor than when utility is linear and there is no discounting, even though the lower of the two cost estimates is used. Here, the role that the oligarchy is willing to play is much reduced.

5.1.5 Discussion

The policy experiments highlight the need for donors of foreign aid to recognize that the oligarchy is wary of losing political control. Policies must be designed accordingly before the oligarchy is willing to share in the costs of education subsidies. For example, if the donor requires that the oligarchy match the subsidy for all poor children, and not just a fraction, the policy will fail. The framework also suggests a policy instrument of education lotteries, to allow the poor to pool income and educate only a subset of themselves. When the gain from the education externality is realized, the remainder of the uneducated poor may be able to educate themselves as well. Finally, even though this is not a model of multiple equilibria with automatic implications for a one-shot policy change, foreign aid is only needed for one generation when utility is linear.

The role of the oligarchy in sharing the costs with foreign donors is diminished when preferences are concave and we allow for discounting. The effect of the education externality, while still an impetus for the rich to subsidize the poor, is weakened. It is also the case

that, with both linear and concave utility, we need to use an implausibly high value for the externality parameter in order to match the model’s constraints. Another concern in this environment is that BV assume that the oligarchy is myopic across generations. They do not care if they lose power, as long as they are not hit with redistributive taxes in *this* generation. The willingness of the oligarchy to subsidize the poor will be weakened, if they are concerned about holding onto power for future generations as well. Lastly, in this model, the oligarchy has complete political control, but BV assume that they do not exercise it by extracting resources from the poor. We next turn to a model that has this feature.

5.2 The Politics of Technology Adoption

5.2.1 Model

The Acemoglu and Robinson (AR) framework consists of a mass of citizens normalized to measure one, an incumbent ruler, and an infinite supply of potential new rulers. Infinitely lived agents discount the future by β . Once rulers are replaced they receive no utility. Each citizen produces $y_t = A_t$, where A_t is the technology available at time t . The ruler taxes output at the rate τ . There is a home production technology that produces $(1 - \tau)A_t$, so that the ruler can only extract taxes at the maximum rate τ from the citizenry. A technological advance increases A by a factor $\alpha > 1$. There is no resource cost to adopting the new technology. However, there may be a cost of political change. The random cost of replacing the ruler when he innovates is $z'A$, and zA if he does not. This cost may be positive or negative; the possibility of a negative draw (a gain) gives citizens an incentive to replace the ruler.

AR assume that $z \sim U[\gamma\mu - 1/2, \gamma\mu + 1/2]$ and $z' \sim U[\mu - 1/2, \mu + 1/2]$, $\gamma \geq 1$. The two key parameters are μ , an inverse measure of the degree of political competition, and γ , a measure of how much technological innovation erodes incumbency advantage by creating “turbulence”. It is optimal for a social planner to always innovate. However, in equilibrium, a ruler will not always innovate if doing so increases the probability that he will be replaced. As γ increases, it is more likely a ruler will not innovate as this decreases the probability of his remaining in power. The effect of μ is more complicated. When $\mu = 0$ (a high degree of political competition), the distributions of z and z' are the same, and a ruler always innovates as γ is now irrelevant. When $\mu \geq 1/2$, the leader is firmly entrenched, is in no danger of losing political power, and therefore innovates.

When $\mu > 0$, and $\gamma > \alpha$, a ruler won’t innovate if $\frac{3}{2}(\alpha - 1)/(\gamma - \alpha) < \mu < \frac{1}{2}$. When $\gamma \leq \alpha$ or $\mu \leq 0$, a ruler will always innovate. This condition is independent of output A , and therefore constant over time.

In the basic model the only benefit to staying in power is a fraction of the output produced. In an extension, AR add another source of extractable rent, R , which is fixed and independent of the prevailing technology. Income at date t is now given by $A_t h$, where h

is an exogenous measure of human capital. Citizens use the same replacement rule as in the basic model, and so the probabilities of staying in power are the same. But now the innovation decision is dependent on A_t . As output grows relative to the fixed rents, the gain from innovating increases. Once an incumbent starts to innovate, he and all future rulers will always innovate. A leader will innovate if and only if $\tau A_t h/R$ is larger than a constant that depends only on the model parameters.⁴⁷ As output grows, this condition will eventually be met. If a country has a higher stock of human capital, or lower rents available to a ruler, innovation is more likely. Interestingly, a higher tax rate, relative to rents, makes innovation more likely, because the ruler gains more from increases in output. Note all inputs are inelastically supplied, and the increased tax rate has no effect on output.

5.2.2 Calibration

We suppose that the “turbulence” parameter γ is structural, while μ is country specific. Consider first countries that always innovate, growing approximately at the frontier growth rate α . While they could be authoritarian or democratic in the model, the countries that we consider are all democratic. The expected length of office of a ruler in a democratic, innovating country is given by the model as, $1/\text{Probability of removal} = 1/(\frac{1}{2} - \mu)$. Data on the tenure of rule identifies the μ associated with each democratic, innovating country. Since these economies are always innovating, it follows from the condition in the previous subsection that $\mu \leq \frac{3}{2}(\alpha - 1)/(\gamma - \alpha)$. We use our collection of frontier countries to pin down an upper bound for γ , γ_U . To do this, we can use the average of μ across all countries, or back out γ from each country using its own μ , and take the upper bound. We can then use γ_U to calculate a $\mu_L = \frac{3}{2}(\alpha - 1)/(\gamma_U - \alpha)$. This procedure essentially allows us to construct an interval $[\mu_L, 1/2]$ within which there is no innovation unless there is a regime change. If $\mu \leq \mu_L$, there is enough political competition and if $\mu \geq 1/2$ there is enough entrenchment, and in both cases there will always be innovation.

We next consider countries where rulers never innovate for fear of losing power. However, an occasional negative cost draw would cause replacement of the ruler, and trigger innovation. The model yields an expression for this minimal growth rate as, $1 + (\alpha - 1)(\alpha - \gamma\mu - 1/2)$. Given α , γ_U , and μ_L , we compute an upper bound on the growth rate that could be rationalized by the model as occurring in politically trapped countries. This, we then compare with the growth rates of sSA countries seen in data to see if we can interpret the sSA situation as a political trap.

We set the average gross annual growth rate of frontier countries, α , to 1.025. Using their average length of rule we calculate $\mu_L = 0.26$.⁴⁸ The procedure outlined above yields

⁴⁷The exact condition can be found in their paper.

⁴⁸In frontier countries – Australia, Canada, Denmark, France, Germany, Italy, Netherlands, Norway, Sweden, the UK, and the US – a leader is in office for an average of 4.24 years (post-World War II period).

$\gamma_U = 1.17$. The annual growth rate has to be less than 0.58% for the model to deliver a no-innovation regime. If we instead compute γ_U as the upper bound resulting from individual country calculations, we get $\gamma_U = 1.13$, and $\mu_L = 0.35$ for Germany. Here, the annual growth rate has to be less than 0.33% for the model to deliver a no-innovation regime.

We approach the extended model the same way. However, data on growth rates, tax rates, and rents are now needed to compute the upper bound, γ_U , and the upper bound on the growth rate consistent with a political trap. While the discount rate β is structural, the rent and tax parameters are country specific. Rewrite $\tau A_t h / R$, which needs to exceed a threshold before innovation happens, as $\tau / (R / A_t h)$, where $A_t h$ is the economy's GDP. We continue to use an α of 1.025, and assume an annual discount rate $\beta = 0.98$. For τ , we use the central government tax revenue as a percentage of GDP for 1990-1995 from the United Nations Online Network in Public Administration and Finance. In the model, taxes literally accrue to the ruler. However, we can use taxes in the data as a proxy for the power a ruler has, and hence for the model's τ . This is 30.8% for the frontier countries and 19.2% for the sSA countries. As a proxy for rents to GDP, we use the percentage of GDP between 1990 and 2002 spent on general administration, also taken from the same data source. Higher administrative expenses are more likely to encourage corruption and rent-seeking activities, which motivated our use of this data. For the frontier countries, this measure is 3.9% and for sSA it is 6.2%. It is the *ratio* of tax rates to rent share that really matters for the innovation condition – this is 7.9 for the frontier countries and 3.1 for the sSA countries. The parameters used are presented in Table 6.

Using the innovating country μ_L of 0.26 that we calculated earlier, we find that $\gamma_U = 1.1615$. Here the annual growth rate has to be less than 0.61% for the model to deliver a no-innovation regime, remarkably close to the 0.58% found in the basic model.

Table 6: Parameters for the Acemoglu-Robinson Models

Parameter	Value	Comment
α	1.025	Average gross annual growth rate of frontier countries
γ	1.13 – 1.17	“Turbulence” parameter; consistent with frontier countries innovating
β	0.98	Annual discount rate
μ	< 0.26, 0.35	Inverse of political competition; consistent with frontier countries innovating
τ	30.8%	Tax revenue as a percentage of GDP for frontier countries, 1990-95; UN data
	19.2%	Tax revenue as a percentage of GDP for sSA countries, 1990-95; UN data
R/Ah	3.9%	% of GDP spent on general administration for frontier countries, 1990-2002; UN data
	6.2%	% of GDP spent on general administration for sSA countries, 1990-2002; UN data

5.2.3 What the Calibrated Models Explain

Using the basic model, can the parameters based on the frontier, innovating countries deliver growth rates consistent with the interpretation that sSA is trapped in a political situation

with little innovation? To answer this, we use the average growth rates for a set of sSA countries from two different sources, WDI and The Penn World Tables (PWT).⁴⁹ If we use the 0.58% bound used from the average μ , 9/15 of the countries satisfy the non-innovation condition when PWT growth rates are used, and 9/11 of the countries satisfy it when the WDI growth rate is used. When the more stringent 0.33% bound is used, 6/15 PWT countries and 7/11 WDI countries are consistent with non-innovation. Therefore, even the basic model is capable of explaining the stagnation in several African countries as arising from a lack of political competition.

In the extended model, even if a country does nothing, it will eventually grow its way out of the trap. How long will that take? If a country grows at a rate of 0.54% per year, close to the average growth rate in sSA, it will take 258 years for the non-innovating condition to be overturned. This is consistent with stagnation in sSA. If a country here reduces its rents to those in developed countries, it will take 172 years for it to become an innovator. If it increases its tax rate to those of developed countries, as well as reduces its rents, it will take 83 years. If it increases its level of democracy to the level of innovating countries, but leaves its rents and taxes at the original levels, it takes 166 years to innovate. Clearly, there is not a big return from piecemeal reforms. To get out of the trap, the democracy level and the tax rate relative to the share of rents would have to be altered simultaneously. This would cause the innovation condition to hold immediately. We are unable to quantitatively evaluate the cost of implementing these recommendations in the way we have done for other models.

However, we can return to the sSA country we have been using as a model of development, Mauritius, for a comparison. Freedom House annually ranks countries on a scale of 1 to 7 on the dimensions of political rights and civil liberties, with 1 being the best and 7 being the worst. The frontier countries we examine have an average score of 1.01 for political rights and 1.25 for civil liberties; that is, they have close to the best ranking. The sSA countries we examine fare much worse, with scores of 5.83 and 5.57. Mauritius, on the other hand has much better scores of 1.77 and 2.23. Likewise, its tax and rent measures are 21.2% and 2.7%, the ratio of which is almost exactly that of the innovating countries. While these facts do not settle issues regarding causality or the exact mechanism posited by the AR model, they are consistent with the model's general premise that the political economy impacts growth.

5.2.4 Discussion

The implied relationship between democracy and growth in this model is not monotonic. High levels of democracy and authoritarianism are both consistent with frontier growth. Therefore, as a country that begins with an authoritarian regime becomes more democratic,

⁴⁹The country list is same as the one in footnote 33, except for Djibouti and Somalia. We use the RGDPCH variable from The Penn World Tables, for an average growth rate of 0.49%, for data available between 1950 and 2000. Using real per capita GNP (1965-1998) from WDI yields an average of -0.15%.

it could first experience a decline in growth before it sees an increase in growth.⁵⁰ The distinction between μ and γ is a useful, and essential feature of the model, but directly pinning down these parameters is a daunting proposition. We have followed an indirect strategy to identify these parameters; therefore, our conclusions should be viewed with caution. Finally, attempting to quantify the costs of altering the degree of democracy is beyond the scope of our paper. In general, a country can also alter the tax rate, rents, or stock of human capital in order to speed the transition to innovation. The aim is to provide a ruler with incentives to innovate. This can obviously be difficult in countries that are undemocratic.

6 A Comparative Evaluation

What conclusions can we draw by considering these models collectively? Consider the ease of calibration first. We have attempted to calibrate models not originally intended for calibration, which poses challenges. Our model, developed explicitly for calibration, was naturally easier in this regard. However, even beyond these obvious differences, the parameters in the non-political models are primarily technological, and are easier to calibrate than those in the more recent political models. With variables such as political rents and democratic turbulence driving explanations for stagnation, there is an element of “theory ahead of measurement” in these models, at the current stage. BV avoid the need for an explicit political variable by equating education with political power, but abstract from the exact mechanism by which this happens. But the political models are the most parsimonious, with the fewest number of parameters to calibrate.

High-quality data is not abundant for the sSA region to which the models are calibrated. Therefore, data used from this region could be considered less reliable than the developed country data used to calibrate the “structural” parameters of the Big Push and occupational choice models. Stagnation in the Big Push model arises from a bad equilibrium selection and except to validate the policy of investment subsidy by comparing it to the one followed in Mauritius, we do not need any sSA data. In the occupational choice model, stagnation arises purely from a bad initial condition, and we need only the income distribution data for our candidate countries of Tanzania and Mauritius. The human capital model on the other hand, relies heavily on sSA data to calibrate education costs relative to income and the skilled-unskilled wage gap. Data from sSA is also needed for the education costs, inequality, political rents, and other variables in the political economy models. However, greater reliance on sSA data increases confidence in the model’s ability to explain the particular situation at hand rather than a generic one.

All the non-political models we consider are capable of generating multiple equilibria. So

⁵⁰This is reminiscent of the non-monotonic connection between democracy and growth that Barro (1996) finds, though he reports democracy and growth are positively related at low levels of democracy (high μ), and negatively at high levels (low μ).

it is natural to ask how robustly this happens for realistic parameter values. In the Big Push model, multiplicity results for a limited range of high fixed costs of industrialization and a high degree of increasing returns. The dynamic investment model is most robust if an open economy assumption is used to sidestep the effect of increased investment on interest rates. A very specific set of parameters is needed in the occupational choice model for multiplicity to result, given tractability considerations. But in both models, the parameters needed are not indefensible from an empirical standpoint. For the human capital model, we followed the strategy of calibrating parameters from *a priori* evidence in our model, and get only the single, stagnant steady state. Therefore, across the models, we conclude that while obtaining multiple equilibria is possible, it cannot be done in a very robust manner. The single-equilibrium political economy models of BV and AR are faced with different robustness considerations. The BV model requires a larger degree of human capital externalities than reported by empirical studies, while we have to resort to debatable proxies such as public administration expenditures for political rents while calibrating the AR model.

Regarding policies to overcome stagnation, the Big Push models suggest investment subsidies and tax credits, the occupational choice model suggests redistribution of initial wealth, and the human capital model suggests education subsidies. The first two are one-shot interventions, while the third is a permanent policy. Given the above discussion on robustness of multiple equilibria, one needs to be cautious about concluding that temporary policies will suffice to revive stagnant economies. While the BV model is not one of multiple equilibria with automatic implications for a one-shot policy change, foreign aid for educating the poor is only needed for one generation.

Quantitatively, the policy interventions suggested by the models are not large: 4 to 7% investment subsidy rate and even lower income tax rate to give the economy a Big Push, 4% of initial wealth redistribution to get a better mix of occupations, around 3% tax rate to subsidize human capital accumulation, and an expenditure of about 1% to 5.6% of lifetime income of the rich in the BV model (provided enough foreign aid is forthcoming for a few years). The AR model provides some insight into why such seemingly low-cost policies are not implemented widely; several political variables such as political rents and democratic power have to be altered first, or in conjunction with other policy changes, and it is not easy to see how this can be accomplished.

Even though the models focus on different channels, considering them together allows us to gain a more comprehensive picture of policy reform. For instance, our human capital policy experiments suggest that foreign aid is not required to revive the fortunes of sSA; by taxing the poor, as opposed to just the rich, which is the only possibility considered in the BV framework, the economy could free itself from stagnation on its own. However, the BV model suggests that temporary foreign aid can act as a seed in encouraging the local wealthy minority to share the costs of educating the poor. Both frameworks favorably view the policy of foreign lending for education purposes, provided problems of enforcement can

be resolved.

We can also identify recurring explanatory factors that would be leading candidates for inclusion in a more comprehensive model. Decreasing inequality would help avoid stagnation in both the occupational choice and the BV setups, and if broadly interpreted, in the human capital framework. Likewise human capital subsidies help in both the BV and AR models, and form the central policy recommendation of the human capital model. Since we associate fixed costs of technology adoption with skilled labor in the calibration of the MSV models, increases in human capital would decrease costs and play a positive role in their setup as well. Finally, capital market imperfections are an important feature of the occupational choice and human capital explanations.

Indeed, Mauritius, the success story of this region has followed several of the suggested policies, such as investment tax credits and human capital subsidies, simultaneously; it also scores highly on surveys of civil rights and political liberties. Since the reasons for stagnation are likely to be multiple, and the cost of each policy not inordinate, following multiple policies instead of a particular one appears to be a prudent approach to follow.

7 Conclusions

Our paper makes both a methodological and a substantive contribution. We use the methodology of calibration to study models of stagnation, each of which captures a different explanation. We think calibration is an ideal choice for evaluating models of stagnation, given the problems of data availability and nonlinearity, and the ease with which it allows the study of counterfactual policy experiments. On the substantive front, we provide quantitative estimates of policy interventions that can overcome economic stagnation. In terms of tax or subsidy rates, or costs as a fraction of GDP, these are not high. However, we see both the need to be cautious about advocating one-shot or temporary policies, and the advantages of following a multipronged approach.

We find that even stylized models can be calibrated with considerable success. And one can find empirical counterparts for the policies they recommend. Given these outcomes, it would be fruitful to extend some of these models with the aim of larger scale computation and calibration. With the burden of analytical tractability reduced, several of the suggested channels of stagnation, including politico-economic factors, could be studied in an integrated fashion in one model, and the costs of the different policy instruments and welfare gains can be compared in a more meaningful way. Finally, it would be useful to collect and assemble more extensive data for use in the quantitative evaluation of models of political economy and stagnation, and focus theoretical attention on how politico-economic reforms could be carried out.

References

- [1] Ablo, Emmanuel, and Ritva Reinikka. 1998. "Do Budgets Really Matter? Evidence From Public Spending on Education and Health in Uganda." The World Bank Policy Research Working Paper 1926.
- [2] Acemoglu, Daron, and Joshua Angrist. 2000. "How Large Are Human Capital Externalities? Evidence from Compulsory Schooling Laws." MIT Working Paper.
- [3] Acemoglu, Daron, Simon Johnson, and James A. Robinson. 2001. "An African Success Story: Botswana." MIT Working Paper.
- [4] Acemoglu, Daron, and James A. Robinson. 2002. "Economic Backwardness in Political Perspective." MIT Working Paper.
- [5] Anker, Richard, Rajendra Paratian, and Raymond Torres. 2001. *Mauritius: Studies on the social dimensions of globalization*. Geneva: International Labour Office.
- [6] Autor, David H., Lawrence F. Katz, and Alan B. Krueger. 1998. "Computing Inequality: Have Computers Changed the Labor Market?" *The Quarterly Journal of Economics*, 113(4), 1169-1213.
- [7] Azariadis, Costas. 1996. "The Economics of Poverty Traps Part One: Complete Markets." *Journal of Economic Growth*, 1, 449-486.
- [8] Azariadis, Costas, and Allan Drazen. "Threshold Externalities in Economic Development." *Quarterly Journal of Economics*, 105(2), 501-526.
- [9] Banerjee, Abhijit V., and Andrew F. Newman. 1995. "Occupational Choice and the Process of Development." *Journal of Political Economy*, 101(2), 274-298.
- [10] Barro, Robert J. 1996. "Democracy and Growth." *Journal of Economic Growth*, 1(1), 1-27.
- [11] Barro, Robert J., and Jong-Wha Lee. 1996. "International Measures of Schooling Years and Schooling Quality." *AEA Papers and Proceedings*, 86(2), 218-223.
- [12] Bartel, Ann P., and Frank R. Lichtenberg. 1987. "The Comparative Advantage of Educated Workers in Implementing New Technology." *The Review of Economics and Statistics*, 69(1), 1-11.
- [13] Becker, Gary S., Kevin M. Murphy, and Robert Tamura. 1990. "Human Capital, Fertility, and Economic Growth." *Journal of Political Economy*, 98(5), S12-37.
- [14] Basu, Susanto, and John G. Fernald. 1997. "Returns to Scale in U.S. Production: Estimates and Implications." *Journal of Political Economy*, 105(2): 249-283.

- [15] Beegle, Kathleen, Rajeev H. Dehejia, and Roberta Gatti. 2002. "Do Households Resort to Child Labor to Cope with Income Shocks?" World Bank Working Paper.
- [16] Bigsten, et. al. 2000. "Rates of Return on Physical and Human Capital in Africa's Manufacturing Sector." *Economic Development and Cultural Change*, 48(4), 801-827.
- [17] Bills, Mark, and Peter J. Klenow. 2000. "Does Schooling Cause Growth?" *American Economic Review*, 90(5), 1160-1183.
- [18] Bourguignon, Francois, and Thierry Verdier 2000. "Oligarchy, democracy, inequality and growth." *Journal of Development Economics*, 62, 285-313.
- [19] Bowles, Samuel, Steven N. Durlauf, and Karla Hoff. 2003. (eds.) *Poverty Traps*, Washington, DC: The World Bank.
- [20] Bredie, Joseph W. B., and Girindre K. Beeharry. 1998. "School Enrollment Decline in Sub-Saharan Africa: Beyond the Supply Constraint." World Bank Discussion Paper No. 395
- [21] Burtless, Gary. 1999. "Risk and Returns of Stock Market Investments Held in Individual Retirement Accounts." *Task Force on Social Security Reform, House Budget Committee, May 11, 1999*. The Brookings Institution.
- [22] Caucutt, Elizabeth M., and Krishna B. Kumar. 2003. "Higher Education Subsidies and Heterogeneity: A Dynamic Analysis." *Journal of Economic Dynamics & Control*, 27(8), 1459-1502.
- [23] Caucutt, Elizabeth M., and Krishna B. Kumar. 2004. "Education Policies to Revive a Stagnant Economy: The Case of sub-Saharan Africa." USC Working Paper.
- [24] De Nardi, Mariacristina, and Marco Cagetti. 2003. "Entrepreneurship, Frictions, and Wealth." *Federal Reserve Bank of Minneapolis, Research Department Staff Report 322*.
- [25] Deininger, Klaus, and Squire, Lyn. 1996. "A New Dataset Measuring Income Inequality." *World Bank Economic Review*, 10, 565-591.
- [26] Dimson, Elroy, Paul Marsh, and Mike Satunton. 2002. *Triumph of the Optimists: 101 Years of Global Investment Returns*. Princeton: Princeton University Press.
- [27] Djankov, Simeon, Rafael La Porta, Florencio Lopez-de-Silanes, and Andrei Shleifer. 2002. "The Regulation of Entry." *The Quarterly Journal of Economics*, 117(1), 1-38.
- [28] Durlauf, Steven N., and Paul Johnson. 1995. "Multiple Regimes and Cross-Country Growth Behavior." *Journal of Applied Econometrics*, 10(4), 365-384.

- [29] Durlauf, Steven N. 1993. "Nonergodic Economic Growth." *Review of Economic Studies*, 60(2), 349-366.
- [30] Galor, Oded, and Joseph Zeira. 1993. "Income Distribution and Macroeconomics." *Review of Economic Studies*, 60, 35-52.
- [31] Greenwood, Jeremy, and Mehmet Yorukoglu. 1997. "1974." *Carnegie-Rochester Conference Series on Public Policy* 46, 49-95.
- [32] Guvenen, Fatih. 2003. "Reconciling Conflicting Evidence on the Elasticity of Intertemporal Substitution: A Macroeconomic Perspective." University of Rochester Working Paper.
- [33] Hall, Robert E. 1988. "The Relation between Price and Marginal Cost in U.S. Industry." *Journal of Political Economy*, 96(5): 921-947.
- [34] Hall, Robert E. 1990. "Invariance Properties of Solow's Productivity Residual," in *Growth, Productivity, Unemployment*, Peter Diamond (ed.), 71-112. Cambridge, MA: MIT Press.
- [35] Hall, Robert E., and Charles I. Jones. 1999. "Why Do Some Countries Produce so Much More Output per Worker than Others?" *The Quarterly Journal of Economics*, 114(1): 83-116.
- [36] Heckman, James J., and Peter J. Klenow. 1997. "Human Capital Policy." University of Chicago Working Paper.
- [37] Lamusse, R. 1995. "Mauritius," in *Exporting Africa: technology, trade and industrialization in Sub-Saharan Africa*, Samuel M. Wangwe (ed.). New York: Routledge.
- [38] Moskowitz, Tobias J., and Annette Vissing-Jorgensen. 2002. "The Returns to Entrepreneurial Investment: A Private Equity Premium Puzzle?" *The American Economic Review*, 92(4), 745-778.
- [39] Murphy, Kevin M., Andrei Shleifer, and Robert W. Vishny. 1989. "Industrialization and the Big Push." *Journal of Political Economy*, 97(5), 1003-1026.
- [40] Ogaki, Masao, and Qiang Zhang. 2001. "Decreasing Relative Risk Aversion and Tests of Risk Sharing." *Econometrica*, 69(2), 515-526.
- [41] Ortin-Ángel, Pedro and Vicente Salas-Fumás. 2002. "Compensation and Span of Control in Hierarchical Organizations." *Journal of Labor Economics*, 20(4), 848-876.
- [42] Parente, Stephen L., and Edward C. Prescott. 1999. "Monopoly Rights: A Barrier to Riches." *American Economic Review*, 89(5), 1216-1233.

- [43] Patterson, Kerry D., and Bahram Pesaran. 1992. "The Intertemporal Elasticity of Consumption in the United States and the United Kingdom." *The Review of Economics and Statistics*, 74(4), 573-584.
- [44] Penrose, Perran. 1998. "Cost Sharing in Education: Public Finance, School and Household Perspectives." Department for International Development Report No. 27.
- [45] Quah, Danny T. 1996. "Convergence Empirics across Economies with (some) Capital Mobility." *Journal of Economic Growth*, 1(1), 95-124.
- [46] Rauch, James E. 1993. "Productivity Gains from Geographical Concentration of Human Capital: Evidence from Cities." *Journal of Urban Economics*, 34, 380-400.
- [47] Stokey, Nancy L. 1998. "Shirtsleeves to Shirtsleeves: The Economics of Social Mobility," in Schwartz, Nancy L., Donald P. Jacobs, Ehud Kalai, and Morton I. Kamien (eds.) *Frontiers of research in economic theory : the Nancy L. Schwartz Memorial Lectures, 1983-1997*. Cambridge, UK: Cambridge University Press.
- [48] Stokey, Nancy L. 1996. "Free Trade, Factor Returns, and Factor Accumulation," *Journal of Economic Growth*, 2(1), 421-447.
- [49] Venniker, Richard. 2000. "Social Returns to Education: A Survey of Recent Literature on Human Capital Externalities." Netherlands Bureau for Economic Policy Analysis, CPB Report 2000/1, 47-50.