Rich nations, poor nations: how much can multiple equilibria explain?[1]

Bryan S. Graham
Department of Economics, Harvard University
Cambridge, MA 02138
bgraham@fas.harvard.edu

Jonathan R. W. Temple
Department of Economics, University of Bristol
8 Woodland Road, Bristol BS8 1TN, United Kingdom
jon.temple@bristol.ac.uk

September 2004

[1]We would like to thank Gavin Cameron, Daniel Cohen, Oded Galor, Michael Kremer, Philip Lane, Omer Moav, Andrew Mountford, Kevin O'Rourke, Stephen Redding, Dani Rodrik and especially Francesco Caselli and Mathan Satchi for helpful comments and discussion. All the usual disclaimers apply. We are also grateful to seminar participants at the universities of Bath, Birkbeck, Bristol, Brunel, Edinburgh, Exeter, Harvard, Manchester, Oxford, Paris, SOAS, Trinity College Dublin, and Warwick, the UCL/LSE growth and development seminar, Ente Einaudi in Rome, the European Winter Meeting of the Econometric Society, the Royal Economic Society annual conference, and the Anglo-French Macro and Finance workshop. Graham would like to thank the Rhodes Trust, the George Webb Medley Endowment Fund, the National Science Foundation, and the Ford Foundation for financial support during various phases of the research leading to this paper. Part of this research was undertaken while both authors were at the University of Oxford. This is a revised version of *CEPR Discussion Paper No. 3046.*

ABSTRACT: This paper asks whether the income gap between rich and poor nations can be explained by multiple equilibria. We explore the quantitative implications of a simple two sector general equilibrium model that gives rise to multiplicity, and calibrate the model for a large number of countries. Under the assumptions of the model, around a quarter of the world's economies are found to be in a low output equilibrium. The output gains associated with an equilibrium switch are sizeable, but well short of the vast income disparity observed in the data.

KEYWORDS: poverty traps, multiple equilibria, TFP differences, calibration

JEL CLASSIFICATION: C00, O14, O41, O47.

1 Introduction

It is common for journalists and popular commentators to present league tables of countries, ranked by GDP per worker. If this was a sports league, the spectators would have long since lost interest: the Spearman rank correlation between GDP per worker in 1970 and GDP per worker in 2000 is 0.91. Or consider the following statistics. At least 35 of the world's countries have never exceeded their 1970 level of GDP per worker by more than 40%. At least 20 have never exceeded their 1970 level by more than 20%.[1]

A popular explanation for these stylized facts, and the vast income differences between rich and poor nations, is the role of poverty traps, or multiple equilibria. The basic idea can be found in Malthus, and has often been revisited by economists interested in the problems of development. In this paper, we extend this line of research by exploring the quantitative properties of a specific model exhibiting multiple equilibria. Our starting point is the two sector, variable-returns-to-scale (VRS) model from the trade theory literature.

We use the VRS model to study the production equilibrium of a small open economy with a moderate extent of increasing returns in non-agriculture (industry and services). This model gives rise to multiplicity for a wide range of parameter values. We then develop and apply a method for calibrating the model when there is more than one equilibrium. Under the assumption that we observe economies in equilibrium, we show how to solve for alternative equilibrium allocations, where they exist. Given the simplicity of the model, the calibration exercise requires only a small number of parameter assumptions, and readily available data.

We take advantage of this simplicity by calibrating the model for 127 countries. Under the maintained assumptions of the model, we can infer whether a given country is in a low output or high output equilibrium. We also examine the conditions on parameters under which a large or small number of the world's countries might be regarded as in a poverty trap. Furthermore, by establishing the nature of any alternative equilibrium solution, we can compute the productivity gains associated with switching from the low output equilibrium to the high output equilibrium. We then compare the current world income distribution with a counterfactual one, where each country is in its high output equilibrium. This allows us to study the extent to which the VRS model can account for the observed variation in levels of income across countries.

The exercise also casts light on international differences in total factor productivity (TFP). In economic environments characterized by multiple equilibria, countries with identical preferences, factor endowments and production technologies may converge to different long run outcomes. This immediately implies that aggregate TFP may differ across countries, even when technologies can be costlessly transferred across national borders. Again, our analysis allows us to quantify this effect. We construct the counterfactual distribution of TFP that would obtain in a world where low output equilibria are

[1] These figures are based on real GDP per worker for 1970-2000 from version 6.1 of the Penn World Tables. The sample excludes countries with populations lower than one million in 2000. The patterns are almost identical when using data on GDP per capita instead of GDP per worker. The figures on stagnation are less striking if 1960 is chosen as the start date, because the 1960s were a good decade for many developing countries. See Easterly (1994) for more on the prevalence of stagnation.

eliminated.

Our main findings are as follows. First, even for a moderate extent of increasing returns to scale, income differences across low output and high output equilibria are sizeable. Output can easily rise by a factor of two or three. But while substantial, these differences are much smaller than the output gaps between the world's richest and poorest economies. The model cannot explain the vast income differences between, say, the USA and sub-Saharan Africa, without assuming implausibly strong increasing returns.

Nevertheless, even for a moderate extent of increasing returns, about a quarter of the world's economies are found to be in a low output equilibrium. The exact proportion varies with parameter choices, but one finding is particularly robust: the model can explain the low incomes of the poorest, predominantly agricultural economies, relative to middle income countries. It performs less well in explaining development levels in middle income East Asia and Latin America, relative to the OECD economies.

The exercise we undertake illuminates some of the strengths and weaknesses of the VRS explanation for underdevelopment, but we should also note the limits to our analysis. We do not provide a formal statistical test of the VRS model. Instead, our aim is to learn more about the quantitative properties of the model, and its potential for explaining certain features of the data under plausible parameter assumptions. Our main focus is whether the multiple equilibria arising in the VRS model could ever give rise to income disparities of the magnitude observed in the data; this could be seen as examining a necessary condition for the VRS version of multiplicity to be a useful modelling device for output differences.

Hence, the paper asks if the model can rationalize the cross-section distribution of GDP per worker at a single point in time. We do not examine whether the model can be used to understand the evolution of the international distribution of GDP per worker over long time spans. Unlike recent theoretical papers such as Galor and Weil (2000), we do not model the dynamics of long-run development, or consider how countries may ultimately emerge from a poverty trap. A closely related limitation to the paper is that we do not model the process of equilibrium selection. In theoretical models in the growth literature, equilibrium selection is often determined by initial conditions, as in the examples considered by Galor (1996). Our paper is closer to ones in which equilibrium selection is determined by expectations, but this does not mean that we see the coordination of expectations as central to the development process. To address this issue would require us to posit a specific model for resolving the indeterminacy in expectations, which could then incorporate many aspects of an economy's history or conditions. Instead, for the purpose of answering our central question, we interpret the multiplicity of equilibria in the VRS model as an outcome associated with some form of coordination failure, and prefer to leave open the precise origins of this coordination failure or the process by which one equilibrium is arrived at rather than another.

The remainder of the paper is structured as follows. The next section discusses the relationship between our work and previous research. Section 3 sets out the VRS model and provides some intuition for the presence of multiple equilibria. Section 4 introduces our calibration strategy. Section 5 describes the data we use and the main parameter assumptions. Sections 6 and 7 present the central results. Section 8 reports the results of a sensitivity analysis, while Section 9 considers the implications of the VRS model for returns to capital. Some technical results and a description of our data can be found in Appendices A to C.

2 Relation to existing literature

Models of multiple equilibria in output levels have a long history in economics. Almost two centuries ago the Reverend Thomas Malthus sketched the possibility in Book II of his *Principles*, observing "that there are many countries not essentially different..., which yet, with nearly equal natural capabilities, make very different progress in wealth" (Malthus 1836:310). In the post-World War II period, the now famous contributions of Rosenstein-Rodan (1943), Nurkse (1954) and Myrdal (1957) aggressively argued that multiple equilibria were crucial for understanding underdevelopment. Even Solow's

(1956) seminal paper on the neoclassical growth model discussed possible extensions generating multiple equilibria.

Recent research, with a particular burst of activity in the late 1980s and early 1990s, has formalized and substantially extended these early contributions. Key papers include those by Azariadis (1996, 2001), Galor (1996), Krugman (1995), Matsuyama (1991), Murphy, Shleifer and Vishny (1989), Rodríguez-Clare (1996) and Rodrik (1995, 1996). More recently, Ray (2000) has emphasized the importance of this general approach to understanding development. Despite the long history of theoretical work, the possibility of multiplicity has rarely influenced empirical research.[2] This is partly because conventional regression methods are not well suited to the analysis of multiple equilibria, and are likely to yield misleading results in this context (see Durlauf and Johnson 1995, Durlauf and Quah 1999 and Galor 1996).[3]

While it is possible to develop formal statistical tests for the presence of multiple equilibria, as in Bloom, Canning and Sevilla (2003), doing so is difficult for a number of reasons. The use of more sophisticated statistical methods often requires larger samples and hence, given the limited number of countries in the world, such techniques can ask a lot of the available data. One concrete manifestation of these challenges is the continuing debate on whether the ergodic world distribution is, or is not, bimodal, with competing statistical approaches generating different answers (e.g., Kremer, Onatski and Stock 2001 versus Quah 2001). For these reasons, we do not pursue an estimation-based strategy in this paper, but instead focus on calibration methods. In particular, we use calibration to explore the quantitative properties of a specific theoretical model. As with other calibration exercises, this should not be seen as a test of the model, but rather as an important step in examining its potential to explain the data. We wish to emphasize our view that the calibration approach pursued here is a useful complement to, and not a substitute for, direct model estimation.

The model we calibrate gives rise to multiplicity in a straightforward way. We consider a small open economy with two sectors, agriculture and non-agriculture. The outputs of both sectors can be traded on world markets. The agricultural production technology has constant returns in labor, capital and a fixed factor that we call land. This implies decreasing returns to the variable inputs. In contrast, the non-agricultural production technology has increasing returns. This model is a specific example of the variable-returns-to-scale (VRS) models of trade theory. It is well known that VRS models can give rise to non-uniqueness, and also have other properties that distinguish them from the more familiar constant returns to scale case.[4] Panagariya (1988) argues that a VRS model can explain a wide variety of stylized facts about the development process.

The introduction of increasing returns creates well known difficulties for modelling. In this paper, following much of the VRS literature, we assume that increasing returns arise through a sector-specific output-generated externality. This greatly simplifies the analysis, because we can retain the standard assumptions of perfect competition and marginal productivity factor pricing.

In the model we calibrate, the presence of multiple equilibria depends crucially upon an externality in the non-agricultural sector. The empirical importance of such externalities remains open to debate. In an influential contribution, Caballero and Lyons (1992) present statistical evidence consistent with substantial externalities across industries. Later work has sometimes cast doubt on the generality of this finding, but one of the contributions of our paper is to show that multiple equilibria can play an important role in generating income differences across countries even when the extent of increasing returns is very modest.[5]

[2] This contrasts sharply with the Solow model, which has served as the organizing framework for influential development accounting and empirical growth research, as in Christensen, Cummings and Jorgenson (1981), Hall and Jones (1999), Klenow and Rodríguez-Clare (1997), Mankiw, Romer and Weil (1992) and Prescott (1998) among many other contributions.

[3] The presence of multiplicity raises distinctive identification issues that are reviewed in detail by Cooper (2001).

[4] The VRS trade literature is reviewed by Choi and Yu (2003). Additional key references include Kemp and Schweinberger (1991) and Panagariya (1981, 1988).

[5] The empircal literature on external economies is briefly surveyed by Benhabib and Farmer (1996, p. 434-435) and Farmer (1999, p. 149-152 and 171-173). Burnside (1996) argues that much of the empirical literature suffers from severe identification problems. More recently, Conley and Dupor (2003) have addressed some of these problems and found that returns to scale are generally constant or modestly increasing, where the precise results depend on the instrument set.

Moreover, broader support for the importance of increasing returns can be found in the cross-country empirical work of Backus *et al* (1992), in the analysis of trade flows by Antweiler and Trefler (2002), and in the work on gravity explanations of trade by Evenett and Keller (2002). The paper by Antweiler and Trefler suggests that increasing returns can be found in about a third of goods-producing industries and form an important source of comparative advantage. Again, their evidence is potentially consistent with the presence of industry-wide externalities, although other effects may also be at work.

Although the econometric evidence for external effects is mixed, there remain strong reasons for exploring the relevance of the VRS model. A number of empirical regularities suggest a role for increasing returns over at least some range of production, notably the well-known tendency for economic activity to be spatially concentrated. External economies or increasing returns can also be useful in explaining business cycles and intra-industry trade flows. As we discuss further below, the VRS model can also account for the "twin peaks" in the cross-country distribution of TFP identified by Feyrer (2003). For all these reasons, we view the VRS model as a natural starting point for a quantitative study of multiplicity.

3 The model

This section describes two versions of a VRS model with multiplicity. The first is a simple one-factor model that we will use to illustrate the principles behind the multiplicity result and our calibration procedure. The second, more complex model is the one that forms the basis for our main calibration results.

In the simple model there is only one factor of production, labor. There are diminishing returns to labor in agriculture and, due to an externality, increasing returns to labor in non-agriculture. We assume that the relative price of the agricultural good is fixed by world prices. We also assume that labor is paid its private marginal product in both sectors (implicitly a fixed factor in agriculture receives the remainder of agricultural income).

The agricultural production function is:

$$Y_a = A_a L_a^{\phi}$$

where $0 < \phi < 1$. For non-agriculture, we introduce a simple form of externality, namely that the output of each individual firm is an increasing function of total non-agricultural employment. Using j to index the firms, we have $Y_{nj} = A_n L_{nj} L_n^{\lambda}$ where Y_{nj} and L_{nj} are the output and employment of firm j and L_n is total employment in this sector. The externality parameter is greater than zero ($\lambda > 0$). Aggregating over firms we then have:

$$Y_n = A_n L_n^{1+\lambda}$$

Firms set employment levels L_{nj} without taking into account their effect on total employment L_n. Hence, the private marginal product and wage in the non-agricultural sector is given by $w_n = A_n L_n^{\lambda}$.

We focus on equilibria with incomplete specialization for which the wage is the same in both sectors. Defining $a = L_a/L$ where $L = L_a + L_n$ is total employment, labor market equilibrium implies:

$$a^{1-\phi}(1-a)^{\lambda} = \phi \left(\frac{A_a}{A_n}\right) L^{\phi-1-\lambda}$$

This equation will usually yield two solutions for the agricultural employment share, a, where $0 < a < 1$. This is the multiplicity result. The underlying intuition is that, due to the externality, the labor demand curve for the non-agricultural sector as a whole is upward sloping: as the sector expands, it can afford to pay higher wages, due to increased productivity.[6] Hence

[6] This is the case even though the demand for labor of individual firms is not upward sloping. Note also that the intuition is similar to that in the one sector business cycle model analysed in Farmer (1999, p.156). The simple one-factor, two-sector model is closely related to models studied by Panagariya (1981) and Ethier (1982b), both of whom provide further intuition for the presence of multiple equilibria.

this upward-sloping curve may intersect more than once with the downward sloping labor demand curve in agriculture.

We now describe the full model that will form the basis for our quantitative results. It is a general equilibrium model of production for a small open economy with two sectors, in which both goods can be traded. As is often assumed in such models, the economy is closed to international flows of capital and labor. Capital and labor are mobile between sectors, while there is a fixed factor in agricultural production, which we call land. The assumption of a sector-specific externality in non-agriculture creates the possibility of non-uniqueness in equilibrium allocations of labor and capital, for reasons that we will discuss further below.[7]

Aggregate output is the sum of agricultural output, Y_a, and non-agricultural output, Y_n:

$$Y = Y_a + Y_n$$

where the assumptions about trade, and appropriate choice of units for the two outputs, allow the prices of both goods to be normalized to one.

We assume that all factors of production are fully employed and that factors receive their private marginal products. We treat the agricultural sector as made up of a large number of landowners and perfectly competitive, profit-maximizing firms. These firms produce using a Cobb-Douglas technology that has constant returns to physical capital, land and labor. Land is rented from the landowners, and all firms pay the same factor costs. Under these assumptions we can restrict attention to a representative firm, and write total agricultural output as:

$$Y_a = [\overline{K} - K_n]^{\alpha} [\overline{R}]^{\beta} [A_a h(\overline{L} - L_n)]^{1-\alpha-\beta} \tag{1}$$

Where $\overline{K}$, $\overline{L}$, and $\overline{R}$ refer respectively to the aggregate (economy-wide) stock of capital, the total labor force, and a fixed quantity of land. For the sake of our later empirical work on productivity differences, we are also allowing the effectiveness of labor to be augmented by an index of human capital, h.[8] Labor-augmenting technology is indexed by A_a.

In the non-agricultural sector, output for firm $j = 1, \ldots, J$ in non-agriculture is given by:

$$\begin{aligned} Y_{nj} &= f(K_{nj}, A_n h L_{nj}) v(Y_n) \\ &= f(k, A_n h) v(Y_n) \, L_{nj} \end{aligned}$$

where $f(.)$ exhibits constant returns to scale and $k \equiv K_n/L_n$ is the capital-labor ratio. The second line follows from the use of constant returns technology by profit-maximizing firms. As is standard, the assumption that all firms within this sector pay the same factor costs, and use the same technology, ensures that all firms choose the same capital-labor ratio.

Note that each firm's output is also a function of total non-agricultural output, Y_n, which could reflect the presence of agglomeration economies or other external effects. In order to retain a standard treatment of competitive equilibrium, we assume that each firm is small enough to disregard its effect on total output. We write total non-agricultural output as:

$$\begin{aligned} Y_n &= \sum Y_{nj} = f(k, A_n h) v(Y_n) \left(\sum L_{nj} \right) \\ &= f(\sum K_{nj}, A_n h \sum L_{nj}) v(Y_n) \\ &= f(K_n, A_n h L_n) v(Y_n) \end{aligned}$$

[7] In the theoretical literature on poverty traps, the presence of external economies arises from explicitly modelled microfoundations, such as a division of labor process, as in Rodriguez-Clare (1996) and Rodrik (1996). Here we concentrate on a simpler model which captures the same fundamental ideas.

[8] From the perspective of calibrating the model for individual countries, our decision to augment labor in both sectors by the same skills factor, h, is equivalent to merely changing the scale of the productivity terms A_a and A_n. Nevertheless, we choose to include an explicit role for human capital at this stage, because cross-country variation in skills will influence our later calculations of international variation in measures of productivity.

For simplicity, we will take $f(.)$ to be Cobb-Douglas, given by $f(.) = K_n^{\gamma}(A_n h L_n)^{1-\gamma}$. We also assume that $v(Y_n) = Y_n^{\frac{\lambda}{1+\lambda}}$. Then output in non-agriculture is

$$Y_n = [K_n]^{\gamma(1+\lambda)}[A_n h L_n]^{(1-\gamma)(1+\lambda)} \tag{2}$$

where the key parameter λ captures the extent of increasing returns for the sector as a whole. For example $\lambda = 0.3$ implies that if all non-agricultural firms simultaneously increase their use of capital and labor by 10 percent, total non-agricultural output would rise by 13 percent (from $1.1^{1.3} \approx 1.13$). The presence of increasing returns means that labor productivity depends on the scale of the sector.

The model is completed by imposing factor market equilibrium conditions. We assume that capital is perfectly mobile between sectors, so that the private returns to capital are equalized, giving[9]

$$r_n = r_a \tag{3}$$

The second equilibrium condition applies to the labor market. Here we leave open the possibility of a sustained wage differential:

$$w_n = w_a(1+\delta) \tag{4}$$

This differential could reflect costs of migration or disutility from urban life, or perhaps some degree of risk aversion together with a lack of informal insurance mechanisms in urban areas.[10] As we will see in calibrating the model, the data appear to imply substantial intersectoral wage differentials under plausible parameter assumptions. Therefore we allow such differentials to be a potential feature of the equilibrium solution of the model.[11]

An equilibrium for our model is defined by an intersectoral allocation of capital and labor such that equations (3) and (4) hold simultaneously.[12] These equilibrium conditions are typically satisfied by more than one allocation of labor and capital, and in the next section we present conditions which ensure two interior equilibria. These different allocations will be associated with different levels of aggregate output per worker and TFP, and also with different factor shares and returns to capital and labor. As in the simpler model, the multiplicity property is driven by the externality in the non-agricultural sector, which yields returns to scale that are increasing for the non-agricultural sector as a whole. Given that capital is mobile, the externality offsets diminishing returns to labor. There will be a range of intersectoral labor allocations for which wages in non-agriculture are increasing in the number of people employed in the sector, even though labor demand curves are always downward sloping at the firm level.

To see this more formally, we first note that if capital mobility ensures that rental rates are equalized, the capital stock employed in the non-agricultural sector, K_n, will be a function of non-agricultural employment L_n and a vector of fixed parameters and constants that we denote by

$$\theta = (\lambda, \gamma, \alpha, \beta, \overline{K}, \overline{D}, \overline{L}, h, A_a, A_n).$$

Hence we can write $K_n = \varphi(L_n, \theta)$ and this implies that the wage in each sector depends on L_n and θ alone. In terms of an equilibrium condition,

$$w_n(L_n; \theta) = (1+\delta) w_a(L_n; \theta). \tag{5}$$

[9]The results are easily generalized to allow for a permanent differential in rental rates across sectors.

[10]Banerjee and Newman (1998) discuss insurance mechanisms in the context of a dual economy. Caselli and Coleman (2001) analyse a model in which a cost of acquiring skills needed for the urban sector leads to a wage differential.

[11]The extent of the differential will have implications for the output gain associated with an equilibrium switch. We explore this issue in our sensitivity analysis later in the paper.

[12]There is another equilibrium, in which all the labor force will be engaged in agricultural production. Since this is never observed in the data, the paper considers only interior equilibria, corresponding to incomplete specialization.

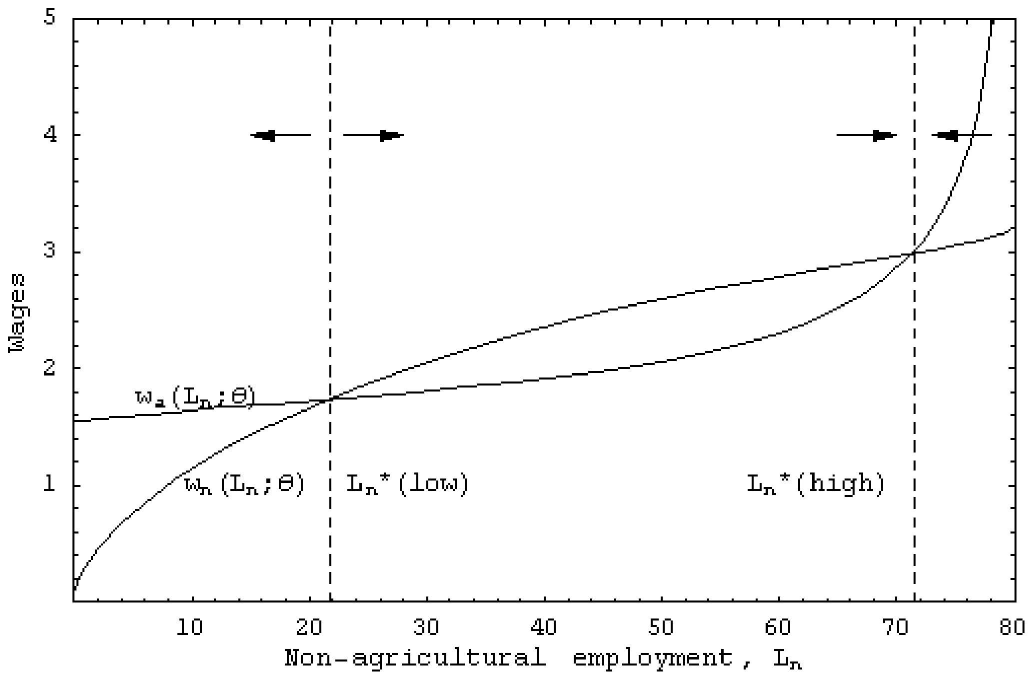

Figure 1: INCREASING RETURNS AND MULTIPLE EQUILIBRIA NOTES: The figure depicts the labor demand curves for non-agriculture, $w_n(L_n;\theta)$, and agriculture, $w_a(L_n;\theta)$, respectively. Equilibria are defined by the intersection of the two curves. The arrows illustrate the dynamics of the system when labor follows a Marshallian adjustment process, where migration between sectors takes place at a rate proportional to the current wage differential.

This reduces our two equilibrium conditions to a single equation in one unknown, namely non-agricultural employment L_n. Figure 1 graphs the right and left-hand-sides of (5) for the case with no equilibrium wage differential ($\delta = 0$). These lines show agricultural and non-agricultural wages as a function of non-agricultural employment, so an equilibrium is defined by the intersection of these two curves. The figure illustrates a case where the curves intersect twice, corresponding to the presence of two interior equilibria.

We end our description of the model with an important digression, on the relevance of dynamics to the VRS model. In calibrating the model, we will be comparing outcomes under alternative static equilibria, and we will not explicitly specify a mechanism for disequilibrium factor adjustment. Consideration of the dynamics remains relevant, however. Depending on the form of the adjustment mechanisms for capital and labor, one of the equilibria may be locally unstable, and thus unlikely to be observed in practice.

Nevertheless, we can establish that local stability for both the interior equilibria is a possible outcome, given a suitably specified intersectoral labor migration process. This is a surprising result since standard approaches to stability analysis - for example positing a Marshallian adjustment process with $\delta = 0$ as depicted in Figure 1 - would imply that the first interior equilibrium is unstable.[13] This is not true, however, provided there is a fixed cost of switching sectors.[14] In this case there will be a range of wage differentials which are insufficient to cover the fixed cost, creating a sphere of stability around both of the equilibria depicted in Figure 1.

We can also show local stability of both interior equilibria in the context of a fully specified migration model, based upon rational and forward-looking decisions. To do this, we build on an insight of Howitt and McAfee (1988). Local stability of

[13] This feature is familiar from related models in the literature, including Panagariya (1981), Diamond (1982), Ethier (1982a, 1982b), Krugman (1991), and Matsuyama (1991).

[14] We are grateful to Michael Kremer for this observation.

the first interior equilibria can be obtained when, in addition to the positive externality emphasized above, a small 'crowding diseconomy' is introduced.[15] The crowding diseconomy means that the cost of intersectoral labor migration is increasing in non-agricultural employment, and this mechanism can produce local stability of all interior equilibria. In the specific example of the appendix, one equilibrium is a sink and the other a saddle point. Since the details of our stability analysis are not essential to grasping the main ideas and concerns of the paper - although essential to making them rigorous - we leave a full treatment to Appendix A.

4 Calibration strategy

This section now describes our strategy for calibrating the model. Throughout, we will use a to denote agriculture's share of employment, as observed in the data, and s to denote agriculture's share of output, again as observed in the data. Note that the output share in both the model and the data corresponds to the nominal share of value added, as evaluated at domestic relative prices. Our calibration method is fully consistent with relative prices that differ across countries, perhaps because of iceberg transport costs, for example.

We assume that we observe the world in equilibrium. More precisely, we assume that the agricultural employment and output shares observed in the data correspond to one of the possible equilibrium allocations. Under this assumption, we can solve for alternative equilibrium allocations in a way that greatly restricts the need for parameter assumptions and data. This makes the calibration exercise unusually straightforward.[16]

Given the observed sectoral structure of employment and output, we use the assumption of intersectoral capital mobility to solve for the sectoral allocations of capital that are compatible with equilibrium. Using the production technologies as well, this allows us to reduce the equilibrium conditions (3) and (4) to a single non-linear equation. Given our assumption that the observed data correspond to one equilibrium solution, this equation depends upon only the technology parameters $(\lambda, \gamma, \alpha, \beta)$ and the agricultural output and employment shares. We can solve this equation numerically to identify all the labor allocations that satisfy the requirements for equilibrium. These solutions will include, by construction, the observed allocation, but also any others consistent with equilibrium.

The great strength of our approach is that it circumvents the need for data on such hard-to-measure variables as the capital stock, total employment and sectoral TFP levels. Initially this may seem counter-intuitive, but we can briefly illustrate the underlying principle by returning to the simple one-factor model described at the start of Section 3. We indicated that, in the simple model, any interior equilibrium allocation of labor (with agricultural employment share a) must satisfy:

$$a^{1-\phi}(1-a)^{\lambda} = \phi\left(\frac{A_a}{A_n}\right)L^{\phi-1-\lambda}$$

Denote the left-hand side as $g(a;\phi,\lambda)$. Note that for an observed value of a and assumed values for ϕ and λ, $g\left(a;\phi,\lambda\right)$ is a known quantity. If there is any other equilibrium labor allocation (which we denote as b) it must satisfy an equation of the same form:

$$b^{1-\phi}(1-b)^{\lambda} = \phi\left(\frac{A_a}{A_n}\right)L^{\phi-1-\lambda}$$

Combining the above two equations means that we can write an equation directly relating the observed employment share,

[15] The interaction of agglomeration economies and crowding diseconomies in the determination of city size is one theme in the new economic geography literature. See Fujita, Krugman and Venables (1999).

[16] Although the assumption that we observe the world in equilibrium may appear a strong one, we have experimented with allowing for simple forms of disequilibrium. These experiments have not modified our general findings.

a, to the alternative equilibrium allocation, b:

$$\begin{aligned} g(b;\phi,\lambda) - g(a;\phi,\lambda) &= 0 \\ &\Rightarrow \left(\frac{a}{b}\right)^{1-\phi}\left(\frac{1-a}{1-b}\right)^{\lambda} - 1 = 0 \end{aligned} \tag{6}$$

Since our assumptions imply that $g(a;\phi,\lambda)$ is a known quantity, we can solve for alternative equilibrium allocations by finding another solution b such that (6) holds. This illustrates that we do not need to know anything about A_a, A_n, or L to derive the alternative equilibrium solution. Instead, all we need to know is the form of the equation $g(.;\phi,\lambda)$ and the value of the employment share found in the data, a.

This principle carries over directly to the more complex model that we calibrate, and to which we now return. We again reduce the equilibrium conditions to one equation and one unknown, where the unknown is the alternative solution for an equilibrium employment share, b, given the observed employment share, a. Appendix B shows that our model implies:

$$\begin{aligned} g(b) - g(a) &= 0 \\ \text{where } g(q) &= q^{\beta}(1-q)^{\lambda}\left[1 - q\left(1 - \frac{\alpha}{\gamma}\frac{s}{1-s}\frac{1-a}{a}\right)\right]^{\alpha-\gamma(1+\lambda)} \end{aligned} \tag{7}$$

Appendix B also shows that, under the parameter restriction

$$\gamma(1+\lambda) - \lambda < \alpha < \gamma\frac{a}{1-a}\frac{1-s}{s} \tag{8}$$

there will almost always be two solutions to (7).[17] One of these solutions will be the observed agricultural employment share, a. The other solution will be the other possible equilibrium allocation of labor, b.

How can we distinguish whether a given country is in a low output or high output equilibrium? The most obvious method is to solve (7) numerically and compare the alternative solution, b, with the observed employment share, a. This is easily done by using Newton's method to solve for the roots of (7). An alternative method for inferring the nature of an equilibrium is available, however. Appendix B shows that it is possible to compare the observed employment share a with a country-specific critical threshold value, a^*, defined as:

$$a^* = \frac{\beta + \frac{\alpha}{\gamma}\frac{s}{1-s}\left[\alpha + \beta - \gamma(1+\lambda)\right]}{\left(1 + \frac{\alpha}{\gamma}\frac{s}{1-s}\right)\left[\alpha + \beta + \lambda - \gamma(1+\lambda)\right]} \tag{9}$$

where a country is in a high output equilibrium when $a < a^*$. Note that for a greater extent of increasing returns (higher λ) the country-specific threshold a^* will be lower for each country. This implies that there will be fewer countries for which the observed agricultural employment share is low enough to indicate a high output equilibrium.

As well as assigning countries to equilibria, our framework allows us to compute the ratio of output in the alternative equilibrium allocation of labor to agriculture (here denoted b) to that in the observed equilibrium (with agricultural employment share a and output share s). Appendix B shows that this ratio is equal to:

$$\Lambda = \frac{Y'}{Y} = \left[\frac{1-b}{1-a}\right]^{(1-\gamma)(1+\lambda)}\left[1 - s\left(\frac{a-b}{a(1-b)}\right)\right] \tag{10}$$

[17] Our parameter choices, to be discussed later, will imply that the restriction is satisfied for almost all of the countries in our data set.

$$\times \left[\frac{1 + \frac{\alpha}{\gamma}\left(\frac{s}{1-s}\right)}{1 + \frac{\alpha}{\gamma}\left(\frac{s}{1-s}\right)\left(\frac{1-a}{a}\right)\left(\frac{b}{1-b}\right)} \right]^{\gamma(1+\lambda)}$$

This equation reveals yet another advantage of the simplicity of the VRS model. We can quantify the output effects of equilibrium switching very easily, using our existing parameter assumptions and data. Again, we do not require knowledge of any of the sectoral TFP parameters or factor endowments.

5 Data and assumptions

In this section, we briefly describe the data and parameter assumptions that will be required to calibrate the VRS model. In assigning each country to a low or high output equilibrium, we require data on agriculture's share of employment and value added (a and s in our notation). These data are taken from the *World Development Indicators CD-ROM* 2000, supplemented with other sources where necessary. We will also use data from Hall and Jones (1999), and data on the stock of agricultural land from the *FAO Yearbook*, when examining international differences in total factor productivity later in the paper. The full data set is described in more detail in Appendix C.

For the three technology parameters, our benchmark case is $\gamma = 0.35$, $\alpha = 0.40$ and $\beta = 0.20$.[18] We have based the value for the non-agricultural capital share, γ, on the aggregate capital share often used in growth accounting (see for example Collins and Bosworth 1996). Our figures for the capital and labor shares in agriculture (α and β) are similar to those used in the GTAP global trade project described in Hertel (1997), which has drawn together data on factor shares from a variety of countries. We have also examined the factor shares implicit in the Martin and Mitra (2001) estimates of a CRS translog production function for agriculture. Their estimates, based on data from the mid-1960s to the present, yield factor shares similar to those adopted here. Later in the paper we will show how our findings on the incidence of poverty traps vary with alternative choices for γ, α and β.

One partial check that our parameter assumptions are plausible is to consider their implications for the aggregate labor share. This share is simply a weighted average of the sectoral labor shares, where the weights are the shares of each sector in nominal value added. Since our production functions are Cobb-Douglas and factors receive their marginal products, the aggregate labor share is:

$$\eta = s(1 - \alpha - \beta) + (1 - s)(1 - \gamma)$$

In our data set, the value of the agricultural output share, s, ranges from 0.3% to 69.3%, with a median of 15.7%. This implies that the aggregate labor share, under our parameter assumptions, will vary between around 48% in the most agricultural countries to around 65% in the least agricultural, with a median of about 61%. Measurement of the aggregate capital share is made difficult, especially in developing countries, by self-employment and unincorporated enterprises (Gollin 2002). Nevertheless, the range of capital shares implicit in our calibration exercise does not look inconsistent with the available data.

The final parameter to consider is λ, which captures the strength of the externality in non-agricultural production. We will report calibration results based on a wide range of values. Caballero and Lyons (1992) find external economies large enough to be consistent with a value for λ of roughly 0.20 to 0.30. Our quantitative exercise will show that even low values for λ, on the order of $\lambda \approx 0.10$, can generate interesting results.

[18] As is standard in the comparative growth literature we assume γ, α, β and λ do not vary across countries. This appears the most sensible approach given the lack of reliable data on sectoral factor shares in developing countries.

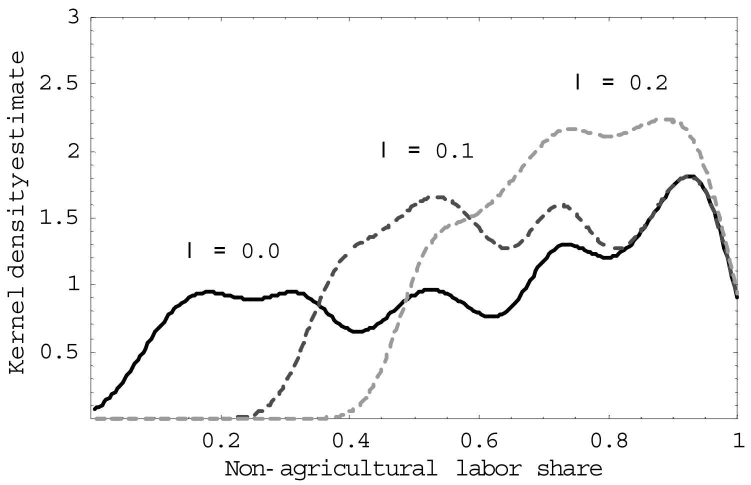

Figure 2: LABOR ALLOCATIONS ACROSS HIGH AND LOW OUTPUT EQUILIBRIA
NOTES: Kernel density estimates for the observed distribution of employment shares across countries, and a selection of counterfactuals, as discussed in the text. Bandwidths were chosen using Sheather and Jones' (1991) plug-in method. The kernel is a univariate standard normal density function.

6 What can the VRS model explain?

We now calibrate the model for 127 countries, as observed in 1988, where the sample is that of Hall and Jones (1999).[19] Under the maintained assumptions of our model, we can derive both of the possible interior equilibrium allocations of labor, where two exist. We can infer the nature of a country's equilibrium, low output or high output, and calculate the ratio of output in the high output equilibrium to that in the low. We can also assess the potential contribution of multiplicity to international variation in output levels, and explore the implications of the model for the world distribution of income.

We begin by presenting the implications of multiplicity for the international pattern of sectoral structure. Recall that, for each country, we observe an equilibrium allocation of employment, while our calibration exercise allows us to compute an alternative equilibrium allocation implied by the model. Figure 2 presents a kernel density plot for the share of employment that is engaged in non-agriculture. The solid line is the distribution observed in the data, while the dotted lines are the counterfactual densities that would obtain if all countries were in their high output equilibrium. (These distributions vary depending on the assumed extent of increasing returns.) As is apparent from the figure, while the actual sectoral distribution of labor is weakly bimodal, the effect of eliminating low output equilibria is to make the distribution of employment patterns look unimodal, so that economies converge in sectoral structure. Note also that the effects of multiplicity are concentrated among the most agricultural economies, an observation that we return to below.

We now address the question posed in the title of our paper, namely, how much of the international variation in output per worker can multiple equilibria explain? To answer this question, we compute summary measures of the cross-country inequality in output per worker. We use inequality measures that are decomposable, so that after dividing our sample of countries into two groups, those in a low and those in a high output equilibrium, we can decompose the sources of inequality into its 'within-group' and 'between-group' components. Note that the assignment of countries to one group or the other, and hence the inequality decomposition, will vary with the assumed extent of increasing returns.

We carry out the inequality decompositions based on two different summary measures of inequality in output per worker. The first measure is a decomposable 'Generalized Entropy' index, due to Shorrocks (1980), with a parameter value of two.

[19] All the calibration experiments were implemented using a program written in Mathematica that is available from the authors on request.

	(1) $I = \frac{1}{2}CV^2 = 0.408$ $\frac{W}{I}(\%)$	$\frac{B}{I}(\%)$	(2) $I = Theil = 0.389$ $\frac{W}{I}(\%)$	$\frac{B}{I}(\%)$	(3) $1 - \frac{Var(\ln y^{HIGH})}{Var(\ln y^{ACTUAL})}$ $(\%)$
$\lambda = 0.05$	81.9	18.1	72.0	28.0	0.146
$\lambda = 0.10$	70.5	29.5	57.5	42.5	0.253
$\lambda = 0.20$	61.4	38.6	49.9	50.1	0.393
$\lambda = 0.30$	53.9	46.1	44.7	55.3	0.481
$\lambda = 0.50$	45.9	54.1	39.9	60.1	0.593

Table 1: MULTIPLICITY AND THE DISPERSION OF INCOME LEVELS

NOTES: Columns (1) and (2) report inequality decompositions. The extent of inequality in output per worker is partitioned into a within-group component ($\frac{W}{I}(\%)$) and a between-group component ($\frac{B}{I}(\%)$). The two groups are countries in low output equilibria and high output equilibria. The decomposition uses a 'Generalized Entropy' measure for inequality with a parameter equal to two (see Mookherjee and Shorrocks 1982) and the Theil coefficient. The former measure is identical to one-half the cofficient of variation squared. Column (3) takes a different approach to quantifying the impact of multiple equilibria on the world distribution of income, and reports one minus the ratio of the counterfactual variance in the log of output per worker (that would be observed if all countries were in their high output equilibrium) to the variance observed in the data.

This measure is equal to one-half of the coefficient of variation squared, and so has a simple relation to the variance. The second measure we use is the Theil index of inequality.

These two decompositions are reported in columns (1) and (2) of Table 1, for five different values of λ. The within-group fraction, corresponding to the proportion of inequality due to the within-group dispersion in levels of income, is denoted $\frac{W}{I}(\%)$. The between-group fraction, corresponding to the proportion of overall inequality due to differences in income levels between those countries in a low output equilibrium and those in a high output one, is denoted $\frac{B}{I}(\%)$. The between-group component provides an upper bound on the extent of the cross-country inequality in output per worker that might be attributed to multiple equilibria.

The results from this exercise are quite dramatic. For the highest values of λ that we examine, we can assign more than half of the inequality in living standards to differences in output levels between those countries in a low output equilibrium and those in a high output one. Even for a value of λ as low as 0.05, our inequality decompositions assign 18 to 28 percent of international inequality to between-group differences. It is essential to note, however, that income levels may vary across low and high equilibrium countries for reasons other than the low or high nature of their current equilibrium. Countries in a low output equilibrium will typically have other characteristics that lead to low output per worker. In this case our decomposition provides only an upper bound on the variation in the cross-country data due to multiple equilibria.

To address this problem, column (3) reports an alternative experiment. First, we compute the variance in the logarithm of output per worker across countries. Next, we force all countries in a low output equilibrium into their alternative, high output equilibrium, calculate the associated increase in output, and then recalculate the variance for this hypothetical or counterfactual distribution. By computing the ratio of the two variances, we can see what would happen to the cross-country variation in living standards if all countries in a low output equilibrium were to switch to their high equilibrium.

Even under this new approach, for a value of λ of 0.20 we can explain around two-fifths of the international variance in log output per worker. Note that in contrast to our first approach, this method probably understates the amount of output variation due to multiple equilibria. This is because it ignores some dynamic or general equilibrium effects that might follow an equilibrium switch, like capital accumulation and increases in schooling.

Next, we ask a closely related question. What would the world distribution of income look like, if all countries were in a high output equilibrium? It has been suggested by Quah (1993, 1997) and Jones (1997) that the world income distribution may be bimodal, or at least tending towards that form. We investigate whether the VRS model can explain this feature. For

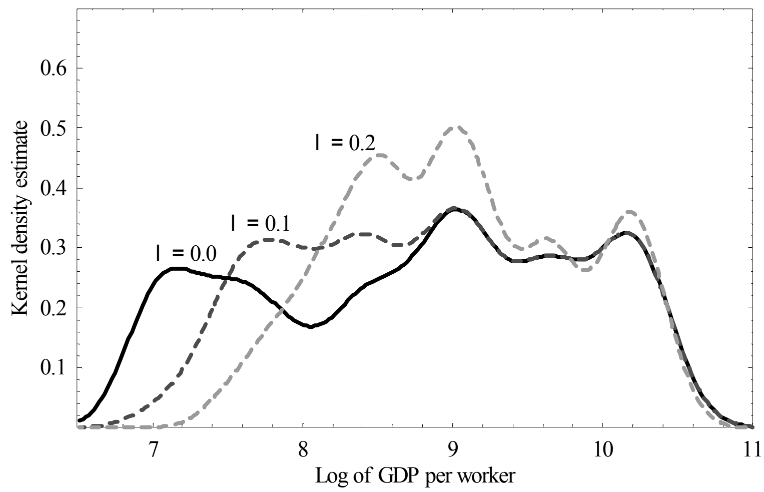

Figure 3: MULTIPLICITY AND THE WORLD DISTRIBUTION OF INCOME (DENSITY ESTIMATES)
NOTES: This figure graphs kernel density estimates for the observed distribution of the logarithm of output per worker across countries, and for the counterfactual distributions that would obtain if all countries were in their high output equilibrium. The cases considered are where λ equals 0.10 and 0.20. Bandwidth selection is by the plug-in method of Sheather and Jones (1991).

simplicity, we do not model the general equilibrium effects that operate at the level of the world as a whole.[20]

Figure 3 plots kernel density estimates for the actual distribution of output per worker (the solid line) and a selection of our counterfactual distributions, generated by placing all countries in their high output equilibrium. Clearly, the poverty traps implied by the VRS model are mainly restricted to the poorest countries. This model appears best suited to explaining income differences between the poorest, predominantly agricultural countries and middle income countries, rather than differences between middle income countries and the OECD.

A remaining question is the size of the output gains associated with a switch to the high output equilibrium. It is clear that, to be of practical interest, a theoretical model of multiplicity should yield equilibria that greatly differ in terms of output per worker. Table 2 offers some insight into this question for our model. We consider a range of cases, corresponding to different values of the externality parameter λ.

We focus on columns (1) and (2) initially. These columns report the number of countries found to be in a low and high output equilibrium respectively. Importantly, even with a moderate extent of increasing returns ($\lambda = 0.10$), we find that roughly a quarter of the countries in our sample are found to be in a low output equilibrium. With a greater extent of increasing returns, the VRS model implies a greater incidence of poverty traps, as indicated previously.

In column (3), we report the mean and median output ratio across the high and low equilibria for countries found to be in a low output equilibrium. We can see that if low output equilibrium economies switched to their high output equilibrium then output per head would approximately double, where the exact figures depend on returns to scale in the two sectors, and the characteristics of the individual countries.[21] The median output gain is similar to the mean, so output gains of this magnitude are not unrepresentative. As noted previously, the output gains are likely to be underestimates, which would be magnified in the presence of endogenous investment in physical capital or schooling.

While these are substantial output gains, they reinforce our conclusion that the model has greatest power in explaining differences between the low end and the middle of the world income distribution. This is consistent with the common idea

[20] The simplification here is to treat world prices as fixed throughout. In principle, we should allow world prices to vary as patterns of sectoral structure change. This is an important consideration if a high proportion of economies are found to be in a low output equilibrium.

[21] There are several influences on this output gain: increasing returns in non-agriculture, decreasing returns in agriculture, and the differential in the marginal product of labor across sectors. We will examine this issue further in Section 8 below.

	(1) # of counties in low (%)	(2) # of counties in high (%)	(3) Mean/Median (y^H/y^L) for "low" group
$\lambda = 0.05$	21 (16.5%)	106 (83.5%)	1.55/1.31
$\lambda = 0.10$	33 (26.0%)	94 (74.0%)	1.70/1.53
$\lambda = 0.20$	44 (34.6%)	83 (65.4%)	2.03/1.70
$\lambda = 0.30$	56 (44.1%)	71 (55.9%)	2.20/1.74
$\lambda = 0.50$	67 (52.8%)	60 (47.2%)	2.69/2.11

Table 2: THE IMPLICATIONS OF MULTIPLE EQUILIBRIA FOR OUTPUT
NOTES: The table reports the distribution of equilibrium assignments for various values of λ. Column (3) reports the mean and median ratio of output in the high output equilibrium to that in the low, for countries found to be in a low output equilibrium.

that poverty trap models are best at illuminating the transition to 'modern economic growth' (in the sense of Kuznets 1966) rather than the differences between middle income countries and the OECD economies.

7 Multiple equilibria and aggregate TFP

In this section, we explore the implications of the VRS model for understanding international differences in total factor productivity (TFP). Several researchers have argued that substantial differences in aggregate TFP are needed to explain the observed variation in output per worker across countries. Some important contributions include those of Hall and Jones (1999), Klenow and Rodriguez-Claré (1997), and Prescott (1998). Subsequent research has focused on developing theoretical explanations for these TFP differences, as in Acemoglu and Zilibotti (2001), Basu and Weil (1998), and Caselli and Coleman (2000).

It is possible that some of the international variation in aggregate TFP reflects the presence of multiple equilibria, rather than technology differences. Consider two countries identical in terms of preferences, technologies and factor endowments, but where one is in a low output equilibrium while the other is in a high output equilibrium. A 'levels accounting' exercise in the tradition of Hall and Jones (1999) will identify a difference in aggregate total factor productivity, even though the sectoral productivity parameters are constant across the two countries.

One reason to study this effect can be found in the work of Feyrer (2003). He finds that the possible bimodality in the world distribution of income is largely explained by bimodality in total factor productivity, rather than in endowments of physical capital or human capital. The VRS model naturally gives rise to bimodality in TFP, whereas other models of multiplicity often imply that bimodality would arise in the distributions of physical or human capital, rather than in TFP.

To investigate this further, we will compare the cross country dispersion of a standard measure of aggregate TFP with the dispersion that would be observed if all countries were in their high output equilibrium.[22] We will show that the VRS model can indeed account for a significant fraction of the international variation in aggregate TFP. For comparison with previous work on TFP differences, we will take into account international differences in human capital. Our adjustments are identical in form to those of Hall and Jones (1999). The basic idea is to construct units of 'effective' labor that are compatible with the Mincerian wage regressions of labor economics. We define $h = e^{\phi(E)}$, where $\phi(E)$ is the efficiency of a worker with E years of schooling relative to a worker with no schooling. The form of $\phi(.)$ is piecewise linear as constructed and parameterized by Hall and Jones (1999). For simplicity, we assume that workers in agriculture and non-agriculture are equally well educated, so effective labor in sector i is then hL_i, where $i = n, a$.

[22] Although our two sector model allows a more sophisticated decomposition of output differences, based on measurement of productivity at the level of the two sectors, here we concentrate on the aggregate decompositions favored by most researchers.

	(1) $V(\log TFP_i - \overline{\log TFP})$	(2) % Variance from Multiplicity	(3) $\bar{\eta}_L$	$\bar{\eta}_D$	$\bar{\eta}_K$
TFP (Actual)	0.234	–	0.600	0.040	0.360
TFP ($\lambda = 0.05$)	0.200	14.5	0.608	0.034	0.358
TFP ($\lambda = 0.10$)	0.177	24.1	0.614	0.029	0.357
TFP ($\lambda = 0.20$)	0.166	29.1	0.622	0.022	0.356
TFP ($\lambda = 0.30$)	0.169	27.6	0.628	0.018	0.354
TFP ($\lambda = 0.50$)	0.203	13.1	0.635	0.012	0.353

Table 3: MULTIPLE EQUILIBRIA AND THE WORLD DISTRIBUTION OF TFP
NOTES: The table lists the variance of the log of aggregate TFP (Column (1)). TFP is calculated using the Christensen, Cummings and Jorgenson (1981) method for the observed data, and then for the counterfactual data obtained by placing all countries in their high output equilibrium. The percentage of the variance due to multiple equilibria, column (2), is then calculated as one minus the ratio of the counterfactual variance of TFP to the observed variance of TFP. Column (3) reports the average factor shares implied by our model for the observed and counterfactual data.

We calculate aggregate TFP using Törnqvist comparisons between individual countries and the world mean, as in Christensen, Cummings and Jorgenson (1981). To make these comparisons, we use the aggregate factor shares implied by our model. Let $\eta_{L,i}$, $\eta_{R,i}$ and $\eta_{K,i}$ be the aggregate income shares for labor, land, and capital in country i. Under the assumptions of our model, it is easily shown that these shares are equal to:

$$\eta_{L,i} = s_i(1-\alpha-\beta) + (1-s_i)(1-\gamma) \tag{11}$$

$$\eta_{R,i} = s_i\beta \tag{12}$$

$$\eta_{K,i} = s_i\alpha + (1-s_i)\gamma \tag{13}$$

where s_i is agriculture's share of output in country i. We then compute the logarithm of aggregate TFP relative to its world mean:

$$\begin{aligned} \log A_i - \overline{\log A} = & \left[\log\frac{Y_i}{L_i} - \overline{\log\frac{Y}{L}}\right] - \left(\frac{\eta_{K,i}+\bar{\eta}_K}{2}\right)\left[\log\frac{K_i}{L_i} - \overline{\log\frac{K}{L}}\right] \\ & -\left(\frac{\eta_{R,i}+\bar{\eta}_R}{2}\right)\left[\log\frac{R_i}{L_i} - \overline{\log\frac{R}{L}}\right] \\ & -\left(\frac{\eta_{L,i}+\bar{\eta}_L}{2}\right)\left[\log h_i - \overline{\log h}\right] \end{aligned} \tag{14}$$

where a line over a variable indicates an average over the countries in our sample.

To obtain the counterfactual distribution of TFP, we force all countries into their high output equilibrium, recalculate (11) to (13) using the new sectoral structure of value added, and then recalculate (14). Table 3 shows the results for various assumptions about the externality parameter, λ. According to the VRS model, up to 30 per cent of the observed variation in the logarithm of aggregate TFP can be explained by multiplicity. Figure 4 shows the cross-country distribution of aggregate TFP and the impact of multiple equilibria on this distribution. The solid line is a kernel density estimate for the observed distribution of the logarithm of TFP. The two dashed lines are density estimates for the counterfactual distributions when $\lambda = 0.10$ and $\lambda = 0.30$ and all countries are in their high output equilibrium. Clearly, allowing for multiplicity leads to a tightening of the aggregate TFP distribution, and moves the distribution towards one that is single-peaked. Yet there is sufficient remaining variation in TFP that the 'costless transfer' view of technology continues to appear problematic, even under the assumptions of the VRS model.

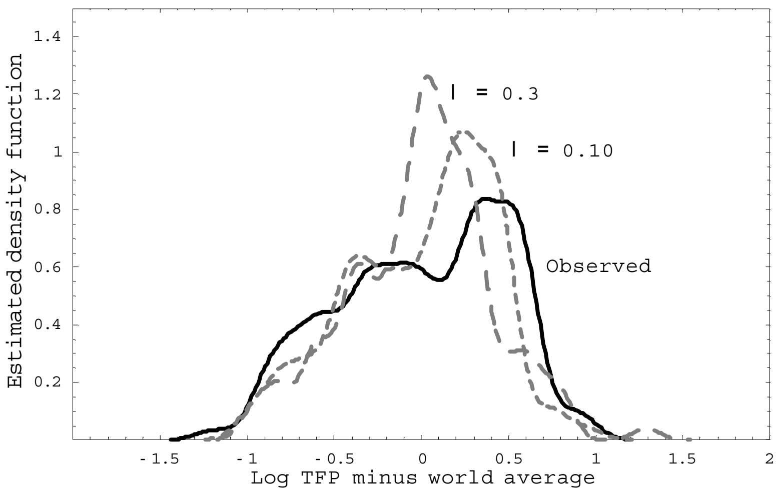

Figure 4: AGGREGATE TFP MEASURES
NOTES: Kernel density estimates for the observed distribution of aggregate TFP across countries and a selection of the counterfactuals discussed in the text. Bandwidths were chosen using Sheather and Jones' (1991) plug-in method. The kernel is a univariate standard normal density function.

Our framework also allows us to investigate productivity measures at the level of sectors, or A_n and A_a in our notation. Given the two sector structure of the VRS model, these are the most appropriate indices of the technologies used by a given country. It is easy to use the techniques in this paper to compute an estimate of TFP for each of the two sectors, and we are exploring this in ongoing research (see Chanda and Dalgaard 2003 for a related contribution).

8 Sensitivity analysis

In this section, we consider whether our main results are sensitive to alternative parameter assumptions. The first part of the section investigates how the incidence of poverty traps will vary with assumptions on the technology parameters α, β, γ and λ. The remainder of the section shows that our main results are not particularly sensitive to reducing the intersectoral differential in the marginal product of labor that is implicit in our previous calculations.

We earlier indicated that the assignment of countries to a low output or high output equilibrium can be achieved by comparing the observed intersectoral labor allocation, a, with a country-specific critical value a^*, given by:

$$a^* = \frac{\beta + \frac{\alpha}{\gamma}\frac{s}{1-s}\left[\alpha + \beta - \gamma(1+\lambda)\right]}{\left(1 + \frac{\alpha}{\gamma}\frac{s}{1-s}\right)\left[\alpha + \beta + \lambda - \gamma(1+\lambda)\right]} \tag{15}$$

where a given country is in a high output equilibrium if $a < a^*$. When the numerator is positive, this critical value for a is decreasing in the magnitude of the externality parameter, λ, and increasing in the exponent on land, β. The greater the extent of increasing returns, the lower is a^*, the fewer economies for which $a < a^*$and so the higher the number of economies found to be in a poverty trap.

By differentiating (15) with respect to α and γ, it is also possible to establish conditions under which a^* is decreasing in α and increasing in γ. These conditions, however, are functions of the three technology parameters (α, β, γ) and also the agricultural employment and output shares, a and s, and so must be evaluated using the data. The conditions are satisfied for almost every single country in our data set when evaluated at our assumed parameter values, but cannot be assumed to

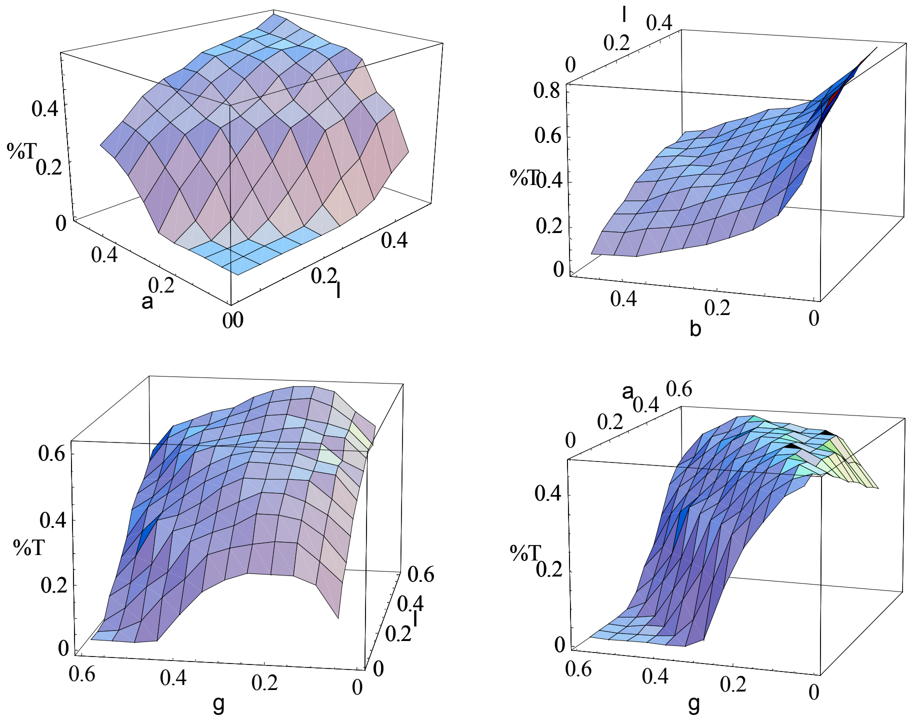

Figure 5: ROBUSTNESS TO DIFFERENT PARAMETER ASSUMPTIONS

NOTES: This figure illustrates the importance of our assumptions regarding α, β, γ and λ for equilibrium assignment. Reading across from the top-left: *Panel 1:* γ and β held fixed at 0.35 and 0.20. *Panel 2:* γ and α held fixed at 0.35 and 0.40. *Panel 3:* α and β held fixed at 0.40 and 0.20. *Panel 4:* β and λ held fixed at 0.20 and 0.30.

hold more broadly.

We therefore investigate the sensitivity of our findings using graphical methods. The four panels of Figure 5 show the fraction of the 127 countries that are found to be in a low output equilibrium, for various configurations of parameters. From the top-right panel, it is clear that the incidence of poverty traps is increasing in λ and decreasing in β, as our theoretical analysis shows. The other panels show the more complex patterns associated with varying the other technology parameters, α, γ and λ, in each case holding one constant and varying the other two.

The second aspect of our sensitivity analysis is more subtle. Recall that we have quantified the output gain associated with an equilibrium switch, for each country in the data set. The size of these output gains depends on the returns to scale in the two sectors, but also on the magnitude of the intersectoral wage gaps. This is because any wage gap corresponds to a marginal product differential. This reinforces the output gain from reallocating labor to non-agriculture, given higher marginal productivity in that sector.

One drawback of our version of the VRS model is that plausible values for technology parameters yield large intersectoral wage differentials.[23] It does seem possible that our previous analysis overstates the extent of marginal product differentials, and therefore overstates the output effects of an equilibrium switch. In the remainder of this section, we investigate the importance of this effect. To do this, we use a simple trick that allows lower wage differentials to be considered while retaining our other assumptions. The trick is to assume that a certain fraction of agricultural output, ϵ, is unmeasured

[23] A likely explanation is that the non-agricultural sector makes more intensive use of skilled labor than agriculture, but allowing for this is not straightforward within the framework of this paper.

	(1) $1 - \frac{Var(\ln y^{HIGH})}{Var(\ln y^{ACTUAL})}$ (%)	(2) % in low	(3) % in high (%)	(4) Mean y^H/y^L for "low" group
$\epsilon = 0.00$	25.3	26.0	74.0	1.70
$\epsilon = 0.10$	24.7	26.0	74.0	1.67
$\epsilon = 0.20$	24.1	26.8	73.2	1.61
$\epsilon = 0.30$	23.3	28.3	71.7	1.54
$\epsilon = 0.40$	22.3	28.3	71.6	1.49
$\epsilon = 0.50$	20.9	30.7	69.3	1.41

Table 4: ROBUSTNESS OF MAIN RESULTS TO MIS-MEASURED AGRICULTURAL OUTPUT
NOTES: Column (1) reports the ratio of the counterfactual variance of the log of GDP per worker, calculated when all countries are in their high output equilibrium, to the observed variance. We consider various assumptions regarding the extent to which agricultural output is undermeasured. Columns (2) and (3) show the consequences for the incidence of poverty traps, while column (4) reveals the consequences for output gains.

in the national accounts. Some agricultural output may be produced for non-marketed household consumption, for example. Under various assumptions about the magnitude of ϵ, we adjust agriculture's share in GDP for this mismeasurement, and then recalculate the equilibrium solutions and the associated welfare effects.

One advantage of this approach is that it has a straightforward effect on the wage differential implicit in our calculations. The Cobb-Douglas production functions imply that marginal products and wages are proportional to average products. Since assuming $\epsilon = 0.5$ implies that 'true' average labor productivity in agriculture is double that in the unadjusted data, it will halve the ratio of wages in non-agriculture to agriculture that is implicit in our calculations.

The adjustments are easily made. We denote true agricultural output as Y_a^* and measured agricultural output as $Y_a = (1-\epsilon)\, Y_a^*$. Hence true output is $Y^* = Y_a^* + pY_n$ and measured output is $Y = Y_a + pY_n$. The data we have on agricultural output shares correspond to $s = Y_a/(Y_a + pY_n)$. We want to calculate the 'true' or adjusted output shares $s^* = Y_a^*/(Y_a^* + pY_n)$. It is easy to show that the 'true' share can be expressed as:

$$s^* = \frac{s}{1 - \epsilon + s\epsilon}$$

Using this result, we adjust the output share data for various possible values of ϵ and then recalculate some of our earlier statistics. The results are shown in Table 4 for the case of $\lambda = 0.10$. The quantity of most interest is the output gain in switching between equilibria. As ϵ is increased from the benchmark case of accurate measurement ($\epsilon = 0$) the effect of an equilibrium switch on output diminishes, consistent with the lower wage ratios that are implicit in these calculations. Yet the reductions in the output gain are relatively modest, implying that our earlier results are not driven simply by large marginal product differentials across sectors.

9 Returns to capital

In this section we briefly consider the implications of the VRS model for returns to capital. The model allows a simple calculation of the ratio of gross marginal products of capital across the two interior equilibria. Given our use of Cobb-Douglas technologies, the ratio of the gross marginal product in the high output equilibrium to that in the low is given by the ratio of the average products of capital, which can be expressed as:

$$\mu_K = \frac{r'}{r} = \left(\frac{s'}{s}\right)\left(\frac{x}{x'}\right)\Lambda$$

	(1) Mean/Median ($\mu_K = r'/r$) for "low" group
$\lambda = 0.05$	1.50/1.26
$\lambda = 0.10$	1.64/1.48
$\lambda = 0.20$	1.95/1.66
$\lambda = 0.30$	2.11/1.68
$\lambda = 0.50$	2.57/1.96

Table 5: RETURNS TO CAPITAL ACROSS EQUILIBRIA
NOTES: Column (1) reports mean and median ratios of the gross marginal product of capital across high and low output equilibria for countries found to be in a low output equilibrium.

where s and x denote the agricultural output share and capital allocation in the low output equilibrium, and s' and x' the corresponding variables for the high output equilibrium. Λ is the ratio of output across the two equilibria, as given by (10). Based on our previous results, Table 5 reports mean and median values for the ratio of gross returns to capital across the two equilibria, for countries found to be in a low output equilibrium. As before, we consider a range of possible values for the externality parameter λ.

It is well known that standard growth models predict a fall in the return to capital as development proceeds, if economies are converging to a steady-state position from below, so that the output-capital ratio is falling. Table 5 indicates that, if growth is partly an outcome of equilibrium switching, the effect of a switch is to offset the decline in the return to capital. This also implies that the VRS model can help in explaining the lack of capital flows to poorer countries, although the size of the effect is too small to be a complete resolution of this puzzle.

10 Conclusion

Although models with multiple equilibria are a popular explanation for the gulf between rich and poor nations, few papers have investigated the empirical relevance of such models. In this paper, we have examined the quantitative implications of a simple model of multiplicity from the VRS class of models in trade theory. We have shown that one such model can be calibrated in a way that greatly limits the need for parameter assumptions and data, and we have exploited this property in order to calibrate the model for 127 countries.

This exercise gives rise to a number of interesting findings. Sections 6 and 8 established that, depending on the choice of parameters, a substantial fraction of countries are found to be in a low output equilibrium, under the maintained assumptions of the model. Moreover, the gains from moving to a high output equilibrium can be sizeable. An immediate corollary is that multiple equilibria can explain a significant fraction of the international inequality in output per worker.

This explanation is by no means a complete one, however. Holding all else constant, a switch to a more productive equilibrium can easily increase output by a factor of two or more, with the precise figure depending on the returns to scale in the two sectors and on the characteristics of individual countries. This is a substantial gain, but falls a long way short of explaining the difference in output per worker between, say, the USA and Niger. The model cannot explain the full extent of the international variation in living standards unless we posit an implausibly strong degree of increasing returns.

The model also implies that only the poorest, predominantly agricultural economies are in a low output equilibrium. Hence, although the VRS model can explain some of the output gap between poor countries and middle income countries, it cannot explain why labor productivity in middle income Latin America and East Asia remains below that of the USA and Western Europe.

The model we have used is a highly stylized one, and the underlying assumptions are strong. Although our calibration results can offer some useful insights into the properties of the VRS model, nothing in this paper approaches a formal test of its validity. Nor do we provide an explanation for why countries arrive at particular equilibria. For all these reasons, we are keen to emphasize that our investigation of multiplicity is far too preliminary to draw any lessons for policy. Nevertheless, we think it is intriguing that even a simple model can give rise to such wide-ranging implications. The quantitative investigation undertaken here has revealed a great deal about the potential explanatory power of one candidate model, and its strengths and limitations. We believe this to be a promising way forward for research on multiple equilibria.

A Stability analysis

This appendix outlines a specification for the dynamics which implies that both interior equilibria are locally stable. It also briefly discusses the approximation involved in our use of a simpler model for calibration purposes.

In specifying the dynamics, we retain the assumption that capital costlessly moves between sectors to equalize rental rates, but now assume that labor migration is costly. In particular, migration costs are described by a convex cost function, $C(\dot{L}_n, L_n)$, where $C_1 > 0, C_{11} > 0, C_2 > 0$ and $C_{12} > 0$ and $C(0, L_n) = 0$. Migration costs are increasing in the current flow of migrants to the non-agricultural sector and in the amount of labor employed in the non-agricultural sector, in both cases due to crowding diseconomies.

The representative individual chooses the level of migration, $\dot{L}_n$, to maximize the present discounted value of income net of migration costs:

$$V(L_n) = \max_{\{\dot{L}_n\}} \int_{\tau=t}^{\infty} e^{-\sigma(\tau-t)} \left[Y_a + Y_n - C(\dot{L}_n, L_n)\right] d\tau \tag{16}$$

where $\sigma$is the discount rate. Since the representative individual is not a social planner she does not take into account the presence of the agglomeration externality in the non-agricultural sector when choosing the optimal migration path. Neither does she internalize the externality generated by the crowding diseconomy in $C(\cdot)$.

We characterize the solution to the sequence problem defined by (16) using a dynamic programming argument. From the envelope theorem, and the fact that the representative individual does not internalize the effects of the agglomeration economies or the crowding diseconomies, we have:

$$\frac{dV(L_n)}{dL_n} = q = \int_{\tau=t}^{\infty} e^{-\sigma(\tau-t)} \left[w_n - w_a\right] d\tau \tag{17}$$

where q is equal to the market value of having an additional unit of labor in non-agriculture instead of agriculture, namely, the expected net present value of future wage premia that an individual realizes from migrating.

To derive a continuous time Bellman Equation we begin with the discrete time optimality condition:

$$V(L_n) = \max_{\{\dot{L}_n\}} \left[(Y_a + Y_n - C(\dot{L}_n, L_n))\Delta t + (1 + \sigma\Delta t)^{-1} V(L_n^{'})\right]$$

where $L_n' = L_n + \Delta L_n$.Rearranging yields:

$$\sigma V(L_n)\Delta t = \max_{\{\dot{L}_n\}} \left[(1 + \sigma\Delta t)(Y_a + Y_n - C(\dot{L}_n, L_n))\Delta t + V(L_n^{'}) - V(L_n)\right]$$

Letting $\Delta t \to 0$ and setting all terms of dt^2or higher equal to zero yields the continuous time Bellman Equation:

$$\sigma V(L_n)dt = \max_{\{\dot{L}_n\}} \left[(Y_a + Y_n - C(\dot{L}_n, L_n))dt + q\dot{L}_n dt\right] \tag{18}$$

The first order condition for optimal migration is:

$$C_1(\dot{L}_n, L_n) = q \tag{19}$$

Migrants therefore equate marginal moving costs with the net present value of expected wage premia from switching sectors. Inverting (19) yields the optimal migration rate:

$$\dot{L}_n = C_1^{-1}(q, L_n) \tag{20}$$

Migration is increasing in the discounted expected stream of future wage premia, q. Optimal behavior, or rationality, also imposes constraints on the path of q over time. Differentiating (17) with respect to time yields:

$$\sigma q = w_n - w_a + \dot{q} \tag{21}$$

The cost of staying in the non-agricultural sector, σq, equals the current wage premium, $w_n - w_a$, plus "capital gains" or increases in the expected value of future wage premia, $\dot{q}$. Equation (21) makes explicit the importance of expectations regarding future wage paths on migration behavior. Positive migration from agriculture to non-agriculture may occur, even if such a movement is associated with an instantaneous wage loss, when the long run gains from switching sectors are expected to be high.

To evaluate the local stability properties of the dynamic system defined by (20) and (21) we linearize the system around a steady state. For simplicity consider the case where $C(\cdot, \cdot)$ takes a quadratic form such that $C(\dot{L}_n, L_n) = 1/2\eta_1(\dot{L}_n)^2 + (1/2\eta_2)(L_n + f)\dot{L}_n$;where f represents fixed moving costs. We can then rewrite (20) as:

$$\dot{L}_n = \eta_1 q - \frac{\eta_1}{\eta_2}(L_n + f) \tag{22}$$

The Jacobian matrix of coefficients for this system is thus:

$$A = \begin{bmatrix} \sigma & -\frac{\delta(w_n - w_a)}{\delta L_n} \\ \eta_1 & -\frac{\eta_1}{\eta_2} \end{bmatrix} \tag{23}$$

The roots of the linearized system $(\varsigma_1, \varsigma_2)$ are defined by $tr(A) = \varsigma_1 + \varsigma_2$ and $\det(A) = \varsigma_1\varsigma_2$. A sufficient condition for local stability of all equilibria is therefore $\sigma < \frac{\eta_1}{\eta_2}$.[24] The interaction of expectations, increasing returns, and a small crowding diseconomy can result in local stability of all interior equilibria if individuals are sufficiently patient (small σ) or the crowding diseconomy is relatively strong (small η_2).

A set of representative dynamics for the system defined by (21) and (22) are depicted in Figure 6. The dynamics of the system can be quite complex, but all our empirical exercise requires is local stability of the two interior equilibria. This is a sufficient condition for both equilibria to be observable under occasional perturbation. For a more detailed discussion of the interaction between expectations and stability in the context of increasing returns, see Howitt and McAfee (1988), Krugman (1991), and Matsuyama (1991). Graham (2000) provides an extensive discussion in the context of the present model.

We now briefly discuss the implications of this analysis for the equilibrium wage gap. The dynamics for $L_n$and q described by (21) and (22) imply that in equilibrium, the intersectoral wage gap will equal

$$w_n - w_a = \frac{\sigma}{\eta_2}(L_n + f). \tag{24}$$

[24]Saddle-stability holds when $det(A) < 0$, which holds at the high output equilibrium, but not the low output equilibrium.

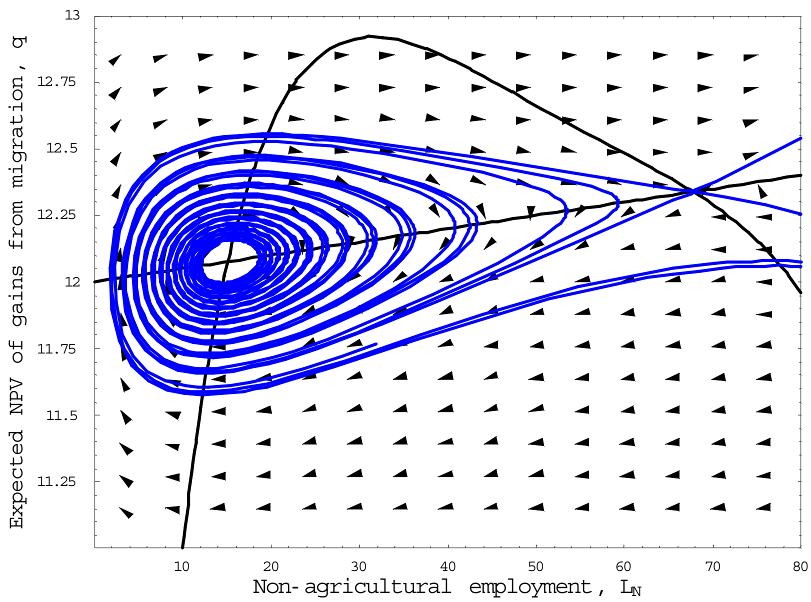

Figure 6: EXPECTATIONS AND EQUILIBRIUM STABILITY
NOTES: The figure depicts the migration dynamics for the system defined by the system of differential equations (22) and (21). The figure illustrates that under our fully specified migration process both of the interior equilibria are locally stable and thus observable under occasional perturbation. However, our migration model also gives rise to local indeterminacy, generating multiple perfect foresight paths for any given initial labor allocation.

In contrast, the model we calibrate is based on a simpler equilibrium condition, namely that the wage ratio is the same for both equilibria. Hence we must think of our calibrated model as only an approximation to the one analyzed here, with its more complete specification of the migration process. In the more complete model, equation (24) shows that the wage gap between sectors is greater in the high output than in the low output equilibrium. This is potentially compatible with our calibration assumption that the wage ratio is fixed across equilibria. Since wages are greater in the high output than in the low output equilibrium, a fixed wage ratio corresponds to a wage gap that is greater in the high output equilibrium.

B Calibration

This appendix shows how to derive an expression that an alternative equilibrium factor allocation must satisfy, and thus forms the basis for our calibration exercise. It also sketches a proof that there will be at most two interior equilibria, under a parameter restriction described below, and provides a simple method for distinguishing whether an observed economy is in a low or high output equilibrium.

In what follows, we denote the fraction of total capital allocated to agriculture in the observed equilibrium by x. $Y_a^{'}$ and $Y_n^{'}$ are the agricultural and non-agricultural output levels associated with the alternative equilibrium, if one exists. The derivation starts from the payment of private marginal products to capital and labor:

$$w_a = (1 - \alpha - \beta)\frac{Y_a}{a\overline{L}} \tag{25}$$

$$r_a = \alpha\frac{Y_a}{x\overline{K}} \tag{26}$$

$$w_n = (1-\gamma)\frac{Y_n}{(1-a)\overline{L}} \tag{27}$$

$$r_n = \gamma\frac{Y_n}{(1-x)\overline{K}} \tag{28}$$

We substitute the expressions for wages into the labor market equilibrium condition (4), which implies:

$$(1-\alpha-\beta)\frac{Y_a}{a\overline{L}} = \left(\frac{1}{1+\delta}\right)(1-\gamma)\frac{Y_n}{(1-a)\overline{L}} \tag{29}$$

This indicates that we can express the ratio of value added in the two sectors as a function of the technology parameters, the agricultural employment share, and the intersectoral wage differential:

$$\frac{Y_a}{Y_n} = \frac{s}{1-s} = \left(\frac{1-\gamma}{1-\alpha-\beta}\right)\left(\frac{a}{1-a}\right)\left(\frac{1}{1+\delta}\right) \tag{30}$$

This equation demonstrates that the intersectoral wage differential $(w_n/w_a = 1+\delta)$ is a function of the technology parameters and the observable variables a and s.

An equation of this form must also hold in any alternative equilibrium with incomplete specialization. For this alternative equilibrium, we use v to denote agriculture's share of value added, and b the share of employment allocated to agriculture. So the corresponding equation will be:

$$\frac{Y_a'}{Y_n'} = \frac{v}{1-v} = \left(\frac{1-\gamma}{1-\alpha-\beta}\right)\left(\frac{b}{1-b}\right)\left(\frac{1}{1+\delta}\right) \tag{31}$$

Equations (30) and (31) imply that:

$$\frac{v}{1-v} = \left(\frac{b}{1-b}\right)\left(\frac{s}{1-s}\right)\left(\frac{1-a}{a}\right) \tag{32}$$

Next we derive an equation for the proportion of capital allocated to agriculture, x, in the observed equilibrium. We denote the rental rate by r $(= r_a = r_n)$. This implies that:

$$x = \frac{K_a}{\bar{K}} = \frac{rK_a}{Y_a}\frac{Y}{r\bar{K}}\frac{Y_a}{Y} \tag{33}$$

$$= \frac{s\alpha}{s\alpha+(1-s)\gamma} \tag{34}$$

where the last line uses marginal productivity factor pricing (so that $rK_a/Y_a = \alpha$, see equation 26) and the identity that holds for the aggregate capital share $(r\bar{K}/Y = s\alpha + (1-s)\gamma)$.

Once again, if there is an alternative equilibrium with incomplete specialization, the fraction of capital allocated to agriculture in that equilibrium (which we denote z) must satisfy a corresponding equation:

$$z = \frac{K_a'}{\bar{K}} = \frac{v\alpha}{v\alpha+(1-v)\gamma}$$

and together with equation (32) this allows us to write z in terms of b, s and a and technology parameters:

$$z = \frac{\frac{b}{1-b}\frac{s}{1-s}\frac{1-a}{a}\left(\frac{\alpha}{\gamma}\right)}{\frac{b}{1-b}\frac{s}{1-s}\frac{1-a}{a}\left(\frac{\alpha}{\gamma}\right)+1} \tag{35}$$

The system of equations is completed by making use of the production functions, to write down the ratio of agricultural to non-agricultural output for both possible equilibria. Using our sectoral production functions (1) and (2) we get:

$$\frac{s}{1-s} = \frac{Y_a}{Y_n} = \frac{[x\overline{K}]^{\alpha} R^{\beta} [A_a h a \overline{L}]^{1-\alpha-\beta}}{[(1-x)\overline{K}]^{\gamma(1+\lambda)} [A_n h (1-a)\overline{L}]^{(1-\gamma)(1+\lambda)}} \tag{36}$$

$$\frac{v}{1-v} = \frac{Y_a'}{Y_n'} = \frac{[z\overline{K}]^{\alpha} R^{\beta} [A_a h b \overline{L}]^{1-\alpha-\beta}}{[(1-z)\overline{K}]^{\gamma(1+\lambda)} [A_n h (1-b)\overline{L}]^{(1-\gamma)(1+\lambda)}} \tag{37}$$

Dividing (37) by (36), simplifying, and using expression (32) yields the following equation:

$$1 = \left(\frac{z}{x}\right)^{\alpha} \left(\frac{1-x}{1-z}\right)^{\gamma(1+\lambda)} \left(\frac{a}{b}\right)^{\alpha+\beta} \left(\frac{1-a}{1-b}\right)^{(1-\gamma)(1+\lambda)-1} \tag{38}$$

The next step is to use the two equations for the intersectoral capital allocations (x and z) to eliminate these variables from (38). When this is done, we obtain our key equation:

$$1 = \left(\frac{a}{b}\right)^{\beta} \left(\frac{1-a}{1-b}\right)^{\lambda} \left[\frac{1-a\left(1-\frac{\alpha}{\gamma}\frac{s}{1-s}\frac{1-a}{a}\right)}{1-b\left(1-\frac{\alpha}{\gamma}\frac{s}{1-s}\frac{1-a}{a}\right)}\right]^{\alpha-\gamma(1+\lambda)} \tag{39}$$

which can be expressed as equation (7) in the text.

This has reduced the system to one equation and one unknown, b. The solutions for $b \in (0,1)$ represent all the interior allocations of labor that satisfy the equilibrium conditions, one of which will be the observed equilibrium labor allocation a and the other an alternative equilibrium allocation, where it exists. The underlying intersectoral capital allocations are easily recovered using equations (33) and (35).

We now investigate the number of equilibria, and also provide a simple way to distinguish whether an observed economy is in a low output or high output equilibrium. To find the equilibrium solutions, we are essentially solving for the roots of the equation in b that is formed by rewriting (39) as:

$$\begin{aligned} g(b) - g(a) &= 0 \\ \text{where } g(q) &= q^{\beta}(1-q)^{\lambda}\left[1-q\left(1-\frac{\alpha}{\gamma}\frac{s}{1-s}\frac{1-a}{a}\right)\right]^{\alpha-\gamma(1+\lambda)} \end{aligned} \tag{40}$$

and $g(a)$ is a known quantity fixed by the data and our parameter assumptions. We now show that under the parameter restriction

$$\gamma(1+\lambda) - \lambda < \alpha < \gamma\frac{a}{1-a}\frac{1-s}{s} \tag{41}$$

there will be at most two equilibria with incomplete specialization.

The underlying idea of the proof can be gained by considering a plot of $\log g(a)$ and $\log g(b)$ against all the potential values of b, that is for $b \in (0,1)$. An example is shown in Figure 7. Since $g(a)$ is a known quantity, $\log g(a)$ is a horizontal line. The intersections of $\log g(b)$ with this horizontal line correspond to the possible equilibrium employment allocations, since at these points $g(a) = g(b)$. Note also that $g(0) = 0$ and $g(1) = 0$.

Hence a sufficient condition for there to exist at most two solutions for b between zero and one is that the curve $\log g(b)$ should be strictly concave. This implies that $\log g(b)$ crosses the straight line $\log g(a)$ twice at most, for values of b where $0 < b < 1$.

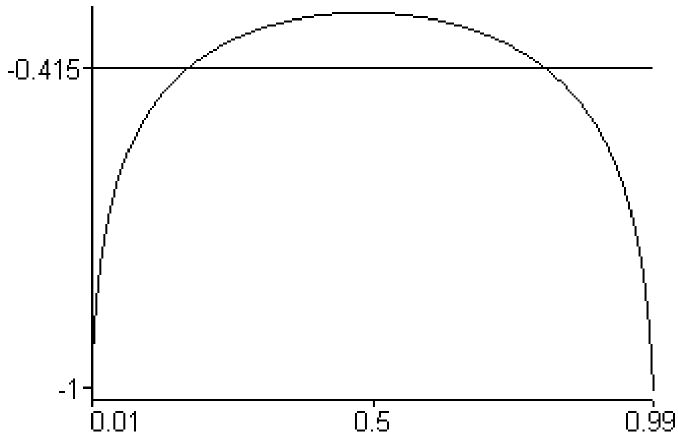

Figure 7: ON THE NUMBER OF EQUILIBRIA
NOTES: An example sketch of log g(a), the horizontal line, and log g(b), the curve. This sketch indicates that there will be at most two interior equilibria under strict concavity of log g(b).

To show strict concavity of $\log g(b)$, we will use (40) to write:

$$g(b) = b^{\beta} (1-b)^{\lambda} [1-bk]^{\theta} \tag{42}$$

where $k = 1 - \frac{\alpha}{\gamma} \frac{s}{1-s} \frac{1-a}{a}$ and $\theta = \alpha - \gamma(1+\lambda)$. Note that our parameter restriction (41) corresponds to assuming $\theta + \lambda > 0$ and $0 < k$. Also, since a, s, α and γ all lie between zero and one, then $k < 1$.

If $\theta > 0$ then it is easy to show that the logarithm of (42) is strictly concave. We now show that this property also holds when $\theta > -\lambda$. We can rewrite (42) as:

$$g(b) = b^{\beta} \left(\frac{1-b}{1-bk} \right)^{\lambda} [1-bk]^{\theta+\lambda}$$

and so:

$$\log g(b) = \beta \log b + \lambda \log \left(\frac{1-b}{1-bk} \right) + (\theta + \lambda) \log(1-bk) \tag{43}$$

Since the sum of strictly concave functions is strictly concave, all we need to show is that each of the three terms in (43) is strictly concave. For the first term, this is obvious by inspection. It is straightforward to show that the final term is also strictly concave (noting that $\theta + \lambda > 0$ by assumption).

The second term can be rewritten as follows:

$$\lambda \log \left(\frac{1-b}{1-bk} \right) = \lambda \log \frac{1}{k} \left(1 - \frac{\frac{1}{k} - 1}{\frac{1}{k} - b} \right)$$

and this implies that the second term is also strictly concave in b, given our assumption that $0 < k < 1$. This completes our sketch of the proof that, under the parameter restriction (41), there will be at most two equilibria with incomplete specialization.

We can also use this result to infer from the observed employment and output shares whether an economy is in its best available equilibrium or one with lower output. Consider the derivative of $g(b)$ with respect to b. The key idea is that for an observed economy to be in its high output equilibrium, and therefore at the leftmost intersection in the figure, this derivative should be greater than zero when evaluated at $b = a$. Conversely if the derivative is less than zero when evaluated at a, the economy must be in a low output equilibrium. There is also a knife-edge case where the derivative of $g(b)$ with respect to b, evaluated at a, is zero. Given strict concavity of $\log g(b)$ this implies that there is only one interior equilibrium.

We can use this idea to generate a condition on the observed employment share (a) that reveals whether or not a given economy's sectoral structure is consistent with a high output equilibrium. The derivative of $g(b)$ with respect to b, evaluated at a, can be written as:

$$\frac{\partial g}{\partial b}|_{b=a} = \left[\frac{\beta}{a} - \frac{\lambda}{1-a} - \frac{\theta k}{1-ak}\right] g(a)$$

Since $ak < 1$ and therefore $g(a) > 0$, the sign of the above expression can be evaluated from the bracketed term alone. Substituting in for θ and k, and rearranging, we can rewrite the necessary condition for a high output equilibrium, namely that $\partial g/\partial b|_{b=a} > 0$ as the condition that $a < a^*$, where the critical value a^* is given by:

$$a^* = \frac{\beta + \frac{\alpha}{\gamma}\frac{s}{1-s}\left[\alpha + \beta - \gamma(1+\lambda)\right]}{\left(1 + \frac{\alpha}{\gamma}\frac{s}{1-s}\right)\left[\alpha + \beta + \lambda - \gamma(1+\lambda)\right]}$$

This is the critical value discussed at several points in the main text.

We now turn to the output effects associated with equilibrium switching, by showing how to calculate the ratio of output in the alternative equilibrium to that in the current one. Here we use b to denote the alternative equilibrium labor allocation to that observed (a). The quantity of interest is:

$$\Lambda_F = \frac{Y_a^{'} + Y_n^{'}}{Y_a + Y_n}$$

which can be rewritten as (recall that units have been chosen to eliminate relative prices):

$$\Lambda_F = \frac{Y_n^{'}\left(1 + \frac{Y_a^{'}}{Y_n^{'}}\right)}{Y_n\left(1 + \frac{Y_a}{Y_n}\right)} \tag{44}$$

Now we make use of the non-agricultural production function, together with equation (32). This yields:

$$\Lambda_F = \left(\frac{1-b}{1-a}\right)^{(1-\gamma)(1+\lambda)} \left(\frac{1-z}{1-x}\right)^{\gamma(1+\lambda)} \left[\frac{1 + \left(\frac{b}{1-b}\right)\left(\frac{1-a}{a}\right)\left(\frac{s}{1-s}\right)}{1 + \frac{s}{1-s}}\right] \tag{45}$$

Substituting in the expressions for z and x and simplifying yields the output ratio associated with an equilibrium switch, equation (10) in the text.

C Data sources

Our data on the sectoral structure of output and employment come primarily from the *World Development Indicators 2000 CD-ROM*. This source is supplemented where necessary with figures drawn from the United Nations' *Yearbook of National Accounts Statistics (1980)* and its successor *National Accounts Statistics: Main Aggregates and Detailed Tables* (1991, 1994, 1995), The United Nations' *Statistical Yearbook CD-ROM*, various *World Development Reports*, and Mitchell's (1998a, 1998b, 1998c) handbooks of historical statistics. In a few cases, the figures were obtained directly from national sources, but over 90 percent of the observations are drawn directly from World Bank or UN sources.

Our dataset for the TFP calculations builds upon that assembled by Hall and Jones (1999). We use their GDP per worker data, as well as their aggregate human and physical capital stock estimates. We add an estimate of agricultural land area using data from the *FAO Yearbook: Production*. Our measure corresponds to arable and pastoral land combined. Since the

Hall and Jones (1999) output data are measured net of mining output, we adjust our sectoral value added data accordingly, using their estimate for the fraction of total GDP due to mining. Data limitations preclude a corresponding adjustment to the employment data, so we assume that total employment in the mining sector is zero. The distortions introduced by this assumption are modest given the enclave nature of mining and its high capital intensity. Even in a highly mineral dependent economy such as Botswana, where roughly one half of GDP is due to mining, less than five percent of the labor force works in the mining sector.

References

[1] Acemoglu, Daron and Fabrizio Zilibotti (2001). "Productivity differences," *Quarterly Journal of Economics 116 (2): 563 - 606.*

[2] Antweiler, Werner and Daniel Trefler (2002). "Increasing returns and all that: a view from trade," *American Economic Review 92 (1): 93 - 119.*

[3] Azariadis, Costas (1996). "The economics of poverty traps part one: complete markets," *Journal of Economic Growth 1 (4): 449 - 486.*

[4] Azariadis, Costas (2001). "The theory of poverty traps: what have we learned?" Manuscript, UCLA.

[5] Backus, David K., Patrick J. Kehoe and Timothy J. Kehoe. (1992). "In search of scale effects in trade and growth," *Journal of Economic Theory 58 (3): 377 - 409.*

[6] Banerjee, Abhijit and Andrew F. Newman. (1998). "Information, the dual economy, and development," *The Review of Economic Studies 65 (4): 631 - 653.*

[7] Basu, Susanto and David N. Weil. (1998). "Appropriate technology and growth," *Quarterly Journal of Economics 113 (4): 1025 - 1054.*

[8] Benhabib, Jess and Roger E. A. Farmer. (1996). "Indeterminacy and sector-specific externalities," *Journal of Monetary Economics 37 (3): 421 - 443.*

[9] Bloom, David E., David Canning, and Jaypee Sevilla. (2003). "Geography and poverty traps," *Journal of Economic Growth, 8(4): 355-378.*

[10] Burnside, Craig. (1996). "Production function regressions, returns to scale, and externalities," *Journal of Monetary Economics 37 (2): 177 - 201.*

[11] Caballero, Ricardo and Richard K. Lyons. (1992). "The case for external economies," *Political Economy, Growth, and Business Cycles* (A. Cukierman *et al*, Eds.). Cambridge, MA: MIT Press.

[12] Caselli, Francesco and Wilbur John Coleman II. (2000). "The world technology frontier," *CEPR Discussion Paper No. 2584.*

[13] Caselli, Francesco and Wilbur John Coleman II. (2001). "The US structural transformation and regional convergence: a reinterpretation," *Journal of Political Economy 109 (3): 584 - 616.*

[14] Chanda, Areendam and Dalgaard, Carl-Johan (2003). "Dual economies and international total factor productivity differences," Manuscript, University of Copenhagen.

[15] Choi, Jai-Young and Yu, Eden S. H. (2003). "External economies in the international trade theory: a survey," *Handbook of International Trade (*E. Kwan Choi & J. Harrigan, Eds.). Oxford: Blackwells.

[16] Christensen, Laurits R., Dianne Cummings, and Dale W. Jorgenson (1981). "Economic growth, 1947 - 1973: an international comparison," *European Economic Review 16 (1): 61 - 94*.

[17] Collins, Susan and Barry Bosworth. (1996). "Economic growth in East Asia: accumulation versus assimilation," *Brookings Papers on Economic Activity 0 (2): 135 - 191.*

[18] Conley, Timothy G. and Bill Dupor (2003). "A spatial analysis of sectoral complementarity," *Journal of Political Economy 111 (2): 311 - 352.*

[19] Cooper, Russell W. (2001). "Estimation and identification of structural parameters in the presence of multiple equilibria," Manuscript, Boston University.

[20] Durlauf, Steven and Danny T. Quah (1999). "The new empirics of economic growth," *Handbook of Macroeconomics* (J. B. Taylor & M. Woodford, Eds.). Amsterdam: North-Holland.

[21] Easterly, William (1994). "Economic stagnation, fixed factors, and policy thresholds," *Journal of Monetary Economics 33 (3): 525 - 557.*

[22] Ethier, Wilfred J. (1982a). "National and international returns to scale in the modern theory of international trade," *American Economic Review 72 (3): 389 - 405.*

[23] Ethier, Wilfred J. (1982b). "Decreasing costs in international trade and Frank Graham's argument for protection," *Econometrica 50 (5): 1243 - 1268.*

[24] Evenett, Simon J. and Keller, Wolfgang (2002). "On theories explaining the success of the gravity equation," *Journal of Political Economy 110 (2): 281 - 316.*

[25] Farmer, Roger E. A. (1999). *The Macroeconomics of Self-fulfilling Prophecies, 2nd Ed.* Cambridge, MA: MIT Press.

[26] Feyrer, James (2003). "Convergence by parts," Manuscript, Dartmouth College.

[27] Food and Agricultural Organization. (1989). *FAO Yearbook: Production.* Rome: Food and Agricultural Organization.

[28] Fujita, Masahisa, Paul Krugman and Anthony J. Venables. (1999). *The Spatial Economy: Cities, Regions, and International Trade.* Cambridge, MA: MIT Press.

[29] Galor, Oded (1996). "Convergence? Inferences from theoretical models," *Economic Journal 106(437): 1056-1069.*

[30] Galor, Oded and Weil, David N. (2000). "Population, technology, and growth: from the Malthusian regime to the demographic transition and beyond," *American Economic Review*, 90, 806-828.

[31] Gollin, Douglas (2002). "Getting income shares right," *Journal of Political Economy 110 (2): 458 - 474.*

[32] Graham, Bryan S. (2000). *Poverty Traps and the World Economy.* M.Phil Thesis, Oxford University.

[33] Hall, Robert E. and Charles I. Jones. (1999). "Why do some countries produce so much more output per worker than others?" *Quarterly Journal of Economics 114 (1): 84 - 116.*

[34] Hertel, Thomas W. (1997). *Global Trade Analysis: Modelling and Applications.* Cambridge: Cambridge University Press.

[35] Howitt, Peter and R. Preston McAfee. (1988). "Stability of equilibria with externalities," *Quarterly Journal of Economics 103 (2): 261 - 277.*

[36] Jones, Charles I. (1997). "On the evolution of the world distribution of income," *Journal of Economic Perspectives 11 (3): 19 - 36.*

[37] Kemp, Murray C. and Schweinberger, Albert G. (1991). Variable returns to scale, non-uniqueness of equilibrium and the gains from international trade. *Review of Economic Studies 58 (4): 807 - 816.*

[38] Klenow, Peter J. and Andrés Rodríguez-Clare. (1997). "The neo-classical revival in growth economics: has it gone too far?" *NBER Macroeconomics Annual 1997* (B.S. Bernanke & J.J. Rotemberg, Eds.), pp. 73 - 114. Cambridge, MA: The MIT Press.

[39] Kremer, Michael, Alexei Onatski and James Stock. (2001). "Searching for prosperity," *Carnegie-Rochester Conference Series on Public Policy 55 (1): 275 - 303.*

[40] Krugman, Paul. (1991). "History versus expectations," *Quarterly Journal of Economics 106 (2): 651 - 667.*

[41] Krugman, Paul. (1995). *Development, Geography, and Economic Theory.* Cambridge, MA: MIT Press.

[42] Kuznets, Simon (1966). *Modern economic growth*. Yale University Press, New Haven.

[43] Malthus, Thomas R. (1836). *Principles of Political Economy: Considered with a View to Their Practical Application, 2nd Ed.* New York: Augustus M. Kelley.

[44] Mankiw, N. Gregory, David Romer and David N. Weil. (1992). "A contribution to the empirics of economic growth," *Quarterly Journal of Economics 107 (2): 407 - 437.*

[45] Martin, Will and Devashish Mitra. (2001). "Productivity growth and convergence in agriculture and manufacturing," *Economic Development and Cultural Change 49 (2): 403 - 422.*

[46] Matsuyama, Kiminori. (1991). "Increasing returns, industrialization and indeterminancy of equilibrium," *Quarterly Journal of Economics 106 (2): 617 - 650.*

[47] Mitchell, Brian R. (1998a). *International Historical Statistics - the Americas, 1750 - 1993.* London: MacMillan Reference.

[48] Mitchell, Brian R. (1998b). *International Historical Statistics - Europe, 1750 - 1993.* London: MacMillan Reference.

[49] Mitchell, Brian R. (1998c). *International Historical Statistics - Africa, Asia, and Oceania, 1750 - 1993.* London: MacMillan Reference.

[50] Mookherjee, Dilip and Anthony F. Shorrocks. (1982). "A decomposition analysis of the trend in UK income inequality," *Economic Journal 92 (368): 886 - 902.*

[51] Murphy, Kevin M., Andrei Shleifer and Robert W. Vishny (1989). "Industrialization and the big push," *Journal of Political Economy 97 (5): 1003 - 1020.*

[52] Myrdal, Gunnar (1957). *Economic Theory and Under-developed Regions.* London: Duckworth.

[53] Nurkse, Ragnar (1953). *Problems of Capital Formation in Underdeveloped Countries.* New York: Oxford University Press.

[54] Panagariya, Arvind. (1981). “Variable returns to scale in production and patterns of specialization,” *American Economic Review 71 (2): 221 - 230.*

[55] Panagariya, Arvind (1988). "A theoretical explanation of some stylized facts of economic growth," *Quarterly Journal of Economics 103 (3): 509 - 526.*

[56] Prescott, Edward C. (1998). “Needed: A theory of total factor productivity,” *International Economic Review 39 (3): 525 - 551.*

[57] Quah, Danny. (1993). “Empirical cross-section dynamics in economic growth,” *European Economic Review 37 (2-3): 426 - 434.*

[58] Quah, Danny (1997). “Empirics for growth and distribution: stratification, polarization, and convergence clubs,” *Journal of Economic Growth 2 (1): 27 - 59.*

[59] Quah, Danny (2001). “Searching for prosperity: a comment,” *Carnegie-Rochester Conference Series on Public Policy 55 (1): 305 - 319.*

[60] Ray, Debraj (2000). “What’s new in development economics?” *The American Economist 44 (2): 3 – 16.*

[61] Rodrik, Dani. (1995). “Getting interventions right: how South Korea and Taiwan grew rich,” *Economic Policy 0 (20): 53 - 97.*

[62] Rodrik, Dani. (1996). “Coordination failure and government policy: a model with applications to East Asia and Eastern Europe," *Journal of International Economics 40 (1-2): 1 - 22.*

[63] Romer, Paul. (1986). “Increasing returns and long-run growth,” *Journal of Political Economy 94 (5): 1002 - 1037.*

[64] Rosenstein-Rodan, P.N. (1943). “Problems with industrialization of Eastern and South-eastern Europe,” *Economic Journal 53 (210/211): 202 - 211.*

[65] Sheather, S.J. and M.C. Jones. (1991). “A reliable data-based bandwidth selection method for kernel density estimation,” *Journal of the Royal Statistical Society B 53 (3): 683 - 690.*

[66] Shorrocks, Anthony F. (1980). “The class of additively decomposable inequality measures,” *Econometrica 48 (3): 613 - 625.*

[67] United Nations, Statistical Office. (1980). *Yearbook of National Accounts Statistics.* New York: United Nations.

[68] United Nations. (1991). *National Accounts Statistics: Main Aggregates and Detailed Tables.* New York: United Nations.

[69] United Nations. (1994). *National Accounts Statistics: Main Aggregates and Detailed Tables.* New York: United Nations.

[70] United Nations. (1995). *National Accounts Statistics: Main Aggregates and Detailed Tables.* New York: United Nations.

[71] United Nations, Statistical Division. (1999). *Statistical Yearbook CD-ROM.* New York: United Nations.

[72] World Bank. (2000). *World Development Indicators 2000 CD-ROM.* Washington D.C.: The World Bank.

D Notes for referees

We first show how to derive equation (38). The first step is to divide (37) by (36). We then have:

$$\frac{v}{1-v}\frac{1-s}{s} = \left(\frac{z}{x}\right)^{\alpha}\left(\frac{1-x}{1-z}\right)^{\gamma(1+\lambda)}\left(\frac{b}{a}\right)^{1-\alpha-\beta}\left(\frac{1-a}{1-b}\right)^{(1-\gamma)(1+\lambda)}$$

Using (32) the left-hand side can be rewritten to give:

$$\frac{b}{1-b}\frac{1-a}{a} = \left(\frac{z}{x}\right)^{\alpha}\left(\frac{1-x}{1-z}\right)^{\gamma(1+\lambda)}\left(\frac{b}{a}\right)^{1-\alpha-\beta}\left(\frac{1-a}{1-b}\right)^{(1-\gamma)(1+\lambda)}$$

which can be simplified to:

$$1 = \left(\frac{z}{x}\right)^{\alpha}\left(\frac{1-x}{1-z}\right)^{\gamma(1+\lambda)}\left(\frac{a}{b}\right)^{\alpha+\beta}\left(\frac{1-a}{1-b}\right)^{(1-\gamma)(1+\lambda)-1} \tag{46}$$

We now need to eliminate x and z. Using (34) and (35) we can write

$$\begin{aligned}\left(\frac{z}{x}\right)^{\alpha} &= \left[\frac{\frac{b}{1-b}\frac{s}{1-s}\frac{1-a}{a}\left(\frac{\alpha}{\gamma}\right)}{\frac{b}{1-b}\frac{s}{1-s}\frac{1-a}{a}\left(\frac{\alpha}{\gamma}\right)+1}\right]^{\alpha}\left(\frac{\frac{s}{1-s}\frac{\alpha}{\gamma}+1}{\frac{s}{1-s}\frac{\alpha}{\gamma}}\right)^{\alpha} \\ &= \left(\frac{b}{1-b}\right)^{\alpha}\left(\frac{1-a}{a}\right)^{\alpha}\left(\frac{\frac{s}{1-s}\frac{\alpha}{\gamma}+1}{\frac{b}{1-b}\frac{s}{1-s}\frac{1-a}{a}\left(\frac{\alpha}{\gamma}\right)+1}\right)^{\alpha}\end{aligned} \tag{47}$$

Similarly we can write:

$$\begin{aligned}\left(\frac{1-x}{1-z}\right)^{\gamma(1+\lambda)} &= \left(\frac{\frac{(1-s)\gamma}{s\alpha+(1-s)\gamma}}{\frac{1}{\frac{b}{1-b}\frac{s}{1-s}\frac{1-a}{a}\left(\frac{\alpha}{\gamma}\right)+1}}\right)^{\gamma(1+\lambda)} \\ &= \left(\frac{\frac{b}{1-b}\frac{s}{1-s}\frac{1-a}{a}\left(\frac{\alpha}{\gamma}\right)+1}{\frac{s}{1-s}\frac{\alpha}{\gamma}+1}\right)^{\gamma(1+\lambda)}\end{aligned} \tag{48}$$

Combining (47) and (48) we get:

$$\left(\frac{z}{x}\right)^{\alpha}\left(\frac{1-x}{1-z}\right)^{\gamma(1+\lambda)} = \left(\frac{b}{1-b}\right)^{\alpha}\left(\frac{1-a}{a}\right)^{\alpha}\left(\frac{\frac{s}{1-s}\frac{\alpha}{\gamma}+1}{\frac{b}{1-b}\frac{s}{1-s}\frac{1-a}{a}\left(\frac{\alpha}{\gamma}\right)+1}\right)^{\alpha-\gamma(1+\lambda)}$$

Substituting this equation into (46) gives:

$$1 = \left(\frac{a}{b}\right)^{\beta}\left(\frac{1-a}{1-b}\right)^{\lambda}\left(\frac{1-a}{1-b}\right)^{\alpha-\gamma(1+\lambda)}\left(\frac{\frac{s}{1-s}\frac{\alpha}{\gamma}+1}{\frac{b}{1-b}\frac{s}{1-s}\frac{1-a}{a}\left(\frac{\alpha}{\gamma}\right)+1}\right)^{\alpha-\gamma(1+\lambda)}$$

Simplifying further:

$$1 = \left(\frac{a}{b}\right)^{\beta}\left(\frac{1-a}{1-b}\right)^{\lambda}\left(\frac{\frac{1-a}{1-b}\frac{s}{1-s}\frac{\alpha}{\gamma}+\frac{1-a}{1-b}}{\frac{b}{1-b}\frac{s}{1-s}\frac{1-a}{a}\left(\frac{\alpha}{\gamma}\right)+1}\right)^{\alpha-\gamma(1+\lambda)}$$

which gives:

$$\begin{aligned}
1 &= \left(\frac{a}{b}\right)^{\beta}\left(\frac{1-a}{1-b}\right)^{\lambda}\left(\frac{(1-a)\frac{s}{1-s}\frac{\alpha}{\gamma}+1-a}{b\frac{s}{1-s}\frac{1-a}{a}\left(\frac{\alpha}{\gamma}\right)+1-b}\right)^{\alpha-\gamma(1+\lambda)} \\
1 &= \left(\frac{a}{b}\right)^{\beta}\left(\frac{1-a}{1-b}\right)^{\lambda}\left(\frac{a(\frac{1-a}{a})\frac{s}{1-s}\frac{\alpha}{\gamma}+1-a}{b\frac{s}{1-s}\frac{1-a}{a}\left(\frac{\alpha}{\gamma}\right)+1-b}\right)^{\alpha-\gamma(1+\lambda)} \\
1 &= \left(\frac{a}{b}\right)^{\beta}\left(\frac{1-a}{1-b}\right)^{\lambda}\left[\frac{1-a\left(1-\frac{\alpha}{\gamma}\frac{s}{1-s}\frac{1-a}{a}\right)}{1-b\left(1-\frac{\alpha}{\gamma}\frac{s}{1-s}\frac{1-a}{a}\right)}\right]^{\alpha-\gamma(1+\lambda)}
\end{aligned}$$

as required.

The next part of these notes fills out the details of the derivation of the a^* criterion, discussed in Appendix B and which appears in sections 4 and 8 of the main text. The starting point is the derivative

$$\frac{\partial g}{\partial b}|_{b=a} = \left[\frac{\beta}{a} - \frac{\lambda}{1-a} - \frac{\theta k}{1-ak}\right] g(a)$$

where $k = 1 - \frac{\alpha}{\gamma}\frac{s}{1-s}\frac{1-a}{a}$ and $\theta = \alpha - \gamma(1+\lambda)$ all as in Appendix B. The condition for an observed economy to be in a high output equilibrium is that the above expression is greater than zero. Since $g(a) > 0$ (see Appendix B) we require:

$$\frac{\beta}{a} - \frac{\lambda}{1-a} - \frac{\theta k}{1-ak} > 0$$

$$\begin{aligned}
\lambda &< (1-a)\left(\frac{\beta}{a} - \frac{\theta k}{1-ak}\right) \\
\lambda &< (1-a)\left(\frac{\beta}{a} - \frac{(\alpha-\gamma(1+\lambda))k)}{1-ak}\right) \\
\lambda &< (1-a)\left(\frac{\beta}{a} - \frac{(\alpha-\gamma-\gamma\lambda)k}{1-ak}\right) \\
\lambda\left[1-\frac{\gamma k(1-a)}{1-ak}\right] &< \left(\frac{1-a}{a}\right)\beta + \frac{(\gamma-\alpha)(1-a)k}{1-ak} \\
\lambda\left[1-ak-\gamma k(1-a)\right] &< \left(\frac{1-a}{a}\right)(1-ak)\beta + (\gamma-\alpha)(1-a)k \\
\lambda\left[1-\gamma k-ak(1-\gamma)\right]a &< (1-a)(1-ak)\beta + (\gamma-\alpha)(1-a)ak
\end{aligned}$$

Thus we arrive at:

$$\lambda\left[1-\gamma k - ak(1-\gamma)\right]\frac{a}{1-a} < \beta - (\alpha+\beta-\gamma)\,ak \tag{49}$$

We now consider the first bracketed term, and substitute in for k:

$$\begin{aligned}1-\gamma k-ak(1-\gamma) &= 1-\gamma+\alpha\frac{s}{1-s}\frac{1-a}{a}-a(1-\gamma)+\frac{\alpha}{\gamma}\frac{s}{1-s}(1-a)(1-\gamma)\\ &= (1-a)\left[1-\gamma+\frac{\alpha}{a}\frac{s}{1-s}+\frac{\alpha}{\gamma}\frac{s}{1-s}(1-\gamma)\right]\\ &= (1-a)\left[1-\gamma+\frac{\alpha}{a}\frac{s}{1-s}\left(1+a\left(\frac{1-\gamma}{\gamma}\right)\right)\right] \qquad (50)\end{aligned}$$

Note that this is greater than zero by inspection.

We next simplify the right-hand-side of (49) again substituting in for k:

$$\begin{aligned}\beta-(\alpha+\beta-\gamma)\,ak &= \beta-a\left(1-\frac{\alpha}{\gamma}\frac{s}{1-s}\frac{1-a}{a}\right)(\alpha+\beta-\gamma)\\ &= \beta-\left(a-\frac{\alpha}{\gamma}\frac{s}{1-s}(1-a)\right)(\alpha+\beta-\gamma)\\ &= \beta-a(\alpha+\beta-\gamma)+\frac{\alpha}{\gamma}\frac{s}{1-s}(1-a)(\alpha+\beta-\gamma)\\ &= \beta(1-a)-a(\alpha-\gamma)+\frac{\alpha}{\gamma}\frac{s}{1-s}(1-a)(\alpha+\beta-\gamma) \qquad (51)\end{aligned}$$

If we use the results in (50) and (51) to rewrite (49) we obtain:

$$a\lambda\left[1-\gamma+\frac{\alpha}{a}\frac{s}{1-s}\left(1+a\left(\frac{1-\gamma}{\gamma}\right)\right)\right]<\beta(1-a)-a(\alpha-\gamma)+\frac{\alpha}{\gamma}\frac{s}{1-s}(1-a)(\alpha+\beta-\gamma)$$

and hence:

$$\lambda<\frac{\beta(1-a)-a(\alpha-\gamma)+\frac{\alpha}{\gamma}\frac{s}{1-s}(1-a)(\alpha+\beta-\gamma)}{a(1-\gamma)+\alpha\frac{s}{1-s}\left(1+a\left(\frac{1-\gamma}{\gamma}\right)\right)}$$

From this equation it is relatively easy to derive inequality (9) in the text, as follows:

$$\lambda<\frac{\beta+a(\gamma-\alpha-\beta)+\frac{\alpha}{\gamma}\frac{s}{1-s}(\gamma-\alpha-\beta)\,a+\frac{\alpha}{\gamma}\frac{s}{1-s}(\alpha+\beta-\gamma)}{a(1-\gamma)\left(1+\frac{\alpha}{\gamma}\frac{s}{1-s}\right)+\alpha\frac{s}{1-s}}$$

which can be rearranged further:

$$\begin{aligned}\lambda\alpha\frac{s}{1-s}+\lambda a(1-\gamma)\left(1+\frac{\alpha}{\gamma}\frac{s}{1-s}\right) &< \beta+a(\gamma-\alpha-\beta)\left[1+\frac{\alpha}{\gamma}\frac{s}{1-s}\right]\\ &\quad+\frac{\alpha}{\gamma}\frac{s}{1-s}(\alpha+\beta-\gamma)\\ a\left(1+\frac{\alpha}{\gamma}\frac{s}{1-s}\right)(\lambda(1-\gamma)-(\gamma-\alpha-\beta)) &< \beta+\frac{\alpha}{\gamma}\frac{s}{1-s}(\alpha+\beta-\gamma)-\lambda\alpha\frac{s}{1-s}\\ a &< \frac{\beta+\frac{\alpha}{\gamma}\frac{s}{1-s}(\alpha+\beta-\gamma-\lambda\gamma)}{\left(1+\frac{\alpha}{\gamma}\frac{s}{1-s}\right)(\lambda(1-\gamma)-(\gamma-\alpha-\beta))}\end{aligned}$$

which can be rearranged to get equation (9) in the text. Note that to preserve the direction of the inequality, the last line requires $\lambda(1-\gamma)-(\gamma-\alpha-\beta)>0$. Our parameter restriction (8) is sufficient to ensure this.

Finally, we show how to derive equation (45). Starting from (44), we have:

$$\Lambda_F = \frac{Y_n^{'}\left(1+\frac{Y_a^{'}}{pY_n^{'}}\right)}{Y_n\left(1+\frac{Y_a}{pY_n}\right)} = \frac{Y_n^{'}\left(1+\frac{v}{1-v}\right)}{Y_n\left(1+\frac{s}{1-s}\right)}$$

Using the production function for non-agriculture, we can rewrite this as:

$$\Lambda_F = \left(\frac{1-b}{1-a}\right)^{(1-\gamma)(1+\lambda)} \left(\frac{1-z}{1-x}\right)^{\gamma(1+\lambda)} \frac{\left(1+\frac{v}{1-v}\right)}{\left(1+\frac{s}{1-s}\right)}$$

Now we can replace the second bracketed term using equation (48) above, to get

$$\Lambda_F = \left(\frac{1-b}{1-a}\right)^{(1-\gamma)(1+\lambda)} \left(\frac{\frac{s}{1-s}\frac{\alpha}{\gamma}+1}{\frac{b}{1-b}\frac{s}{1-s}\frac{1-a}{a}\left(\frac{\alpha}{\gamma}\right)+1}\right)^{\gamma(1+\lambda)} \left(\frac{1+\frac{v}{1-v}}{1+\frac{s}{1-s}}\right)$$

We now focus on eliminating v from the third bracketed term. We can do this using equation (32). This gives

$$\begin{aligned}
\Lambda_F &= \left(\frac{1-b}{1-a}\right)^{(1-\gamma)(1+\lambda)} \left(\frac{\frac{s}{1-s}\frac{\alpha}{\gamma}+1}{\frac{b}{1-b}\frac{s}{1-s}\frac{1-a}{a}\left(\frac{\alpha}{\gamma}\right)+1}\right)^{\gamma(1+\lambda)} \left[\frac{1+\left(\frac{b}{1-b}\right)\left(\frac{s}{1-s}\right)\left(\frac{1-a}{a}\right)}{\left(1+\frac{s}{1-s}\right)}\right] \\
&= \left(\frac{1-b}{1-a}\right)^{(1-\gamma)(1+\lambda)} \left(\frac{\frac{s}{1-s}\frac{\alpha}{\gamma}+1}{\frac{b}{1-b}\frac{s}{1-s}\frac{1-a}{a}\left(\frac{\alpha}{\gamma}\right)+1}\right)^{\gamma(1+\lambda)} \left[1-s+s\left(\frac{b}{1-b}\right)\left(\frac{1-a}{a}\right)\right] \\
&= \left(\frac{1-b}{1-a}\right)^{(1-\gamma)(1+\lambda)} \left(\frac{\frac{s}{1-s}\frac{\alpha}{\gamma}+1}{\frac{b}{1-b}\frac{s}{1-s}\frac{1-a}{a}\left(\frac{\alpha}{\gamma}\right)+1}\right)^{\gamma(1+\lambda)} \left[1-s\left(1-\left(\frac{b}{1-b}\right)\left(\frac{1-a}{a}\right)\right)\right] \\
&= \left(\frac{1-b}{1-a}\right)^{(1-\gamma)(1+\lambda)} \left(\frac{\frac{s}{1-s}\frac{\alpha}{\gamma}+1}{\frac{b}{1-b}\frac{s}{1-s}\frac{1-a}{a}\left(\frac{\alpha}{\gamma}\right)+1}\right)^{\gamma(1+\lambda)} \left[1-s\left(\frac{a(1-b)-b(1-a)}{a(1-b)}\right)\right] \\
&= \left(\frac{1-b}{1-a}\right)^{(1-\gamma)(1+\lambda)} \left(\frac{\frac{s}{1-s}\frac{\alpha}{\gamma}+1}{\frac{b}{1-b}\frac{s}{1-s}\frac{1-a}{a}\left(\frac{\alpha}{\gamma}\right)+1}\right)^{\gamma(1+\lambda)} \left[1-s\left(\frac{a-b}{a(1-b)}\right)\right]
\end{aligned}$$

which is a simple rearrangement of equation (10) in the text.