Abstract

We apply the Business Cycle Accounting methodology developed by Chari, Kehoe and McGrattan (2007) to study the economic resurgence of Brazil, Russia, India and China (BRIC) over the last decade. We document that while efficiency wedges do contribute in a large part to growth, especially in Brazil and Russia, there is an increasing importance of investment wedges especially in the late 2000s, noted in China and India. The results are typically related to the stages of development with Brazil and Russia coming off a crisis to grow in the 2000s, while India and China were already on a stable growth path. Our conclusions are robust to alternative measurements of wedges as well as model extensions allowing investment adjustment costs. Relating wedge patterns to institutional and financial reforms, we find that financial market developments and effective governance in BRICs in the last decade are consistent with improvements in investment and efficiency wedges that led to growth.

JEL Codes: E32

Keywords: BRIC, business cycle accounting, efficiency, market frictions, trend shocks, investment adjustment costs

*We thank Vincenzo Quadrini, Robert Dekle, Guillarme Vandenbroucke, Joel David, Jagjit Chadha, Miguel Leon-Ledesman and participants at the University of Southern California Applied Economics seminar for helpful comments and Tetsuaki Takano for excellent research assistance. All remaining errors are our own.

†Dept. of Economics, University of San Francisco, 2130 Fulton Street | San Francisco, CA 94117-1080; Tel: +1 415 422-4715; Email: schakraborty2@usfca.edu

‡School of Economics, University of Kent, Canterbury, Kent, CT2 7NP, United Kingdom, Tel: +44 1227-827305; Email: K.Otsu@kent.ac.uk.
At its simplest, a growth economy should be regarded as one that is likely to experience rising productivity, which, together with favorable demographics, points to economic growth that outpaces the global average......So we opted for the following: any economy outside the so-called developed world that accounts for at least 1% of current global GDP should be defined as a growth economy. — — — — — – Jim O’Neill (M.D. & Head of Global Economic Research at Goldman Sachs)

1 Introduction

Over the last decade, the average growth rate of the quartet, Brazil, Russia, India and China (known by the acronym BRIC coined by O’Neill in 2001) has outpaced the global average. Cumulative share of the BRIC nations in the world gross domestic product (GDP) has grown from about 16% in 2000 to 26% in 2011 earning China and India the second and third spots in world GDP rankings (the top spot still belongs to the United States), with Russia and Brazil taking the sixth and the seventh spots (Table 1). The trade volume of the group currently takes up 15% of the world trade and jointly, this group of countries is home to about 40% of the world population.

The broad facts of BRIC growth are generally well known. In Table 2, we compare the growth rates in aggregate and per capita GDP in the BRIC nations with that of the United States and the OECD since 1960s. A few interesting facts emerge. While Brazil and India started the 1960s closer to their US and OECD counterparts, China faltered. During the 1970s, while China played catch-up and Brazil continued its economic growth, Indian growth started to decline. The tables turned in the 1980s with Brazilian growth slowing as India made a come-back. China continued on its path of economic growth. 1990s were a period of turbulence with Brazil unable to recover from the 1980s lost decade and the newly formed Russian Federation (1991) facing recession with the Russian financial crisis in 1998. India too began the 1990s with financial trouble with the the possibility of defaulting on its loans and with practically depleted foreign exchange reserves, while China faced political unrest and economic uncertainty due to the Asian financial crisis. However, growth numbers from the 1990s suggest that while Brazil and Russia stagnated, the economic performance of India and China remained relatively stable in the face of

1Tables 1 and 2 are from the IMF and Angus Maddison’s online data resources
2Per capita GDP growth rate in Brazil was low as compared to the aggregate GDP growth due to high population growth.
economic and political troubles. Finally, during the last decade of 2000s, all BRIC nations made a remarkable come-back, with China leading the pack with double-digit economic growth.

The purpose of this paper is to analyze the fluctuations in output growth of the BRIC economies during 1990 to 2009 using a Business Cycle Accounting (BCA) “wedge” methodology formulated by Cole and Ohanian (2004) and Chari, Kehoe and McGrattan (henceforth CKM, 2007) amongst others.

The BCA methodology allows us to quantitatively account for the role played by changes in productivity and factor market distortions in generating output fluctuations by applying a two-pronged approach. BCA uses a real business cycle framework to model various frictions as "wedges" that keep the economy from achieving a first best outcome. These wedges show up as distortions in the first order conditions. Efficiency wedges appear as time-varying productivity. Labor and investment wedges appear as “taxes” on labor and capital income, where “taxes” represent broadly the distortions affecting the labor and investment decisions. Government consumption wedge appears as government expenditure (in a closed economy setup, net exports are also added to government expenditure). In step one, the first order conditions of the model along with data on output, consumption, investment and labor are used to estimate the wedges. In step two, the estimated wedges from step 1 are fed back into the model individually and in different combinations to ascertain their marginal contributions in generating the observed economic outcome. These wedges are the “channels” through which external forces like institutional or policy changes affect the economy.

Comparing the remarkable performance of the BRICs in the last decade with that of the earlier decade of the 1990s, we identify two distinct mechanisms at work: i) in Brazil and Russia, that emerged from a crisis in the 1990s to experience sharp growth in the 2000s, distortions in the investment and labor market (particularly in Brazil) are responsible for the relative stagnation during the 1990s while improvement in production efficiency is the single most important factor in accounting for the rapid growth in the 2000s; ii) in contrast, in India and China which were on a relatively stable growth path since the 1990s, while changes in production efficiency account for a large part of the output fluctuations over the two decades, decline in the investment market distortions become increasingly important in the 2000s, particularly accounting for growth in the latter half. In none of the economies do labor wedges play any role in accounting for growth in the 2000s. Government consumption wedges partially aids China\(^3\) but is ineffective in the other three nations.

\(^3\)The role of government consumption wedges turn out to be model specific. While it plays a minimal role in our benchmark, its contribution increases in the alternative models considered.
However, as we discuss in later sections, this does not mean that government policies are unimportant. What our BCA results tell us is that whatever policy or institutional changes (the "primary drivers") were responsible for the rapid growth of the 2000s worked primarily by increasing production efficiency or by reducing investment market frictions. This finding is particularly interesting as existing BCA literature finds little impact of investment frictions on output during sharp recession periods, attributing most business cycle fluctuations to changes in productivity. For our set of countries examined here, we find that investment wedges are important in accounting for the decade long slowdown in Brazil and Russia in the 1990s and the growth in India and China in the 2000s through gradual capital accumulation.

Our findings are robust to two checks we conduct. Firstly, our benchmark model (in the tradition of BCA literature) assumes that efficiency wedges are transitory fluctuations of productivity about its "trend" to which the economy eventually returns. For our first robustness check, we consider efficiency wedges as shocks to the trend of productivity in the spirit of Aguiar and Gopinath (2007). How we define efficiency wedges matters for investment wedges as well since the latter depends on the expectations about future efficiency. As expected, alternative definition of efficiency wedge affects the measurement of efficiency and investment wedges, however, we essentially find that the roles played by them are similar to those in the benchmark case. As a second robustness check, we add capital adjustment costs assuming that it is technologically costly to convert output into installed capital. As argued by Christiano and Davis (2006), the model simulations with investment wedges is sensitive to inclusion or non-inclusion of investment adjustment costs and can non-trivially affect the conclusions, however in our case, we find that our primary conclusions do not change.

Our accounting work can be related to two distinct strands of literature. Literature on BRIC nations have primarily focused on isolating the singular causes of growth, primarily focusing on India and China (Song, Storesletten and Zilibotti, 2011; Dekle and Vandenbroucke, 2011; Fujiwara, Otsu and Saito, 2011; Bosworth and Collins, 2008; Jones and Sahu, 2009). Focus of Brazil and Russia has been primarily to explain their business cycle downturns primarily in the late 1980s and 1990s (Braginsky and Myerson, 2007; Merlevede, Schoors and Aarle, 2007; Kanczuk, 2004). What distinguishes our study from these previous strands of research is that while most of the earlier literature focuses on the primary drivers of growth, our focus is on identifying the channels through which these external drivers work to stimulate the economy. Secondly our study is related to the extensive literature applying BCA to study economic fluctuations (CKM, 2007; Graminha 2006; Kersting, 2004).

---

4Aguiar and Gopinath (2007) simulate a model with both transitory and trend shocks and find that emerging economies are often characterized by shocks to the trend component.
While most existing literature applies BCA to understand crisis, analysis of growth is sparse, with the exception being Lu (2012). Our study adds to the existing BCA literature by studying BRIC growth through the lens of BCA.

Our accounting results so far suggest an important role of efficiency and investment wedges in the BRIC economies. In our final section, we attempt to tie the observed wedge patterns with some indices of institutional and policy changes in the BRICs. A growing literature in recent years have found microlevel evidence of influence of credit market movements on investment and economic growth both across nations as well as in emerging economies (Bekaert, Harvey and Lundblad, 2011; Alfaro, Kalemsli-Ozcan and Sayek, 2009). Consistent with the earlier literature, we observe an improvement in credit worthiness as well as access to credit in all the BRIC nations that is consistent with declining investment market frictions and increasing efficiency. In addition, while not all institutional and governance indicators that we examine are consistent with observed improvements in efficiency and investment climate, improvements in political stability to some extent since mid-2000s (particularly in Russia) and government effectiveness to a large degree are consistent with observed time series patterns of efficiency and investment wedges. However, the BRICs still have a long distance to go to catch up to the developed West in other areas of governance like control of corruption or rule of law.

The remainder of the paper is organized as follows. In section 2 we describe the business cycle accounting model. In section 3 we explain the business cycle accounting procedure and present the results. In section 4 we provide sensitivity analysis results. In section 5 we discuss the underlying factors that can explain the evolution of wedges. Section 6 concludes the paper.

## 2 The Model

Traditional BCA methodology relies on a standard, closed economy RBC model with a representative household, firm and a government. The representative firm hires labor and capital from the household to produce output using a constant returns to scale technology, which is affected by time-varying production efficiency. The representative household decides on consumption, labor and investment each period. The

---

5Chakraborty (2010), Ljungwall and Gao (2009) and Hsu and Zhao (2009) are some additional studies to focus on growth, but mainly in India and China in isolated time periods. To the best of our knowledge, our paper is the first to conduct a BCA analysis for the Russian economy.
household faces a budget constraint where its expenditure is limited by its labor and capital income. In addition, as the ultimate owner of the firm, the consumer receives the profits. The consumer pays distortionary taxes on labor and capital income to the government. In the BCA framework, these distortionary taxes represent broader economic distortions that affect the factor markets. The government uses its tax revenue to finance government consumption. Any remaining amount is transferred back to the households as lump sum transfers. Exogenous shocks to production efficiency, government consumption and distortionary tax rates are revealed in the beginning of each period and affect economic incentives.

2.1 Firm

The representative firm borrows capital \( K_t \) and labor \( L_t \) from the household in order to produce output \( Y_t \) according to a Cobb-Douglas production function:

\[
Y_t = K_t^\theta (A_t L_t)^{1-\theta},
\]

where \( A_t \) denotes exogenous production efficiency. Labor is defined as total hours worked (product of employment and hours worked per worker). Productivity can be divided into a trend component \( \Gamma_t \) and a cyclical component \( \gamma_t \), i.e. \( A_t = \gamma_t \Gamma_t \), where we assume a constant growth rate in the trend component:

\[
\frac{\Gamma_t}{\Gamma_{t-1}} = a.
\]

Labor grows over time due to growth in population \( N_t \) where we assume a constant growth rate in population:

\[
\frac{N_t}{N_{t-1}} = n.
\]

Output and capital grows over time due to both population and productivity growth. All variables are detrended by the growth trends in order to define a stationary problem:

\[
y_t = \frac{Y_t}{N_t \Gamma_t}, k_t = \frac{K_t}{N_t \Gamma_t}, l_t = \frac{L_t}{N_t}, \gamma_t = \frac{A_t}{\Gamma_t}.
\]

Firms maximize profits \( \pi_t \):

\[
\max \pi_t = y_t - r_t k_t - w_t l_t
\]

where \( r_t \) and \( w_t \) denote the real return on capital and the real wage respectively. The detrended production function can be rewritten as

\[
y_t = k_t^\theta (\gamma_t l_t)^{1-\theta}.
\]
For the benchmark model, we follow CKM (2007) and define the efficiency wedges as

\[ \omega_{e,t} = \gamma_t. \]  

(3)

2.2 The Household and Government

The representative household gains utility from consumption \( c_t \) and leisure \( 1 - l_t \) where we assume a log-linear utility function for our analysis:

\[ u(c_t, 1 - l_t) = \Psi \ln c_t + (1 - \Psi) \ln(1 - l_t). \]

Total hours available is normalized to one\(^6\). The household maximizes its expected lifetime utility:

\[ \max E_t \sum \beta^t [u(c_t, 1 - l_t)], \]

where \( \beta \) is the subjective discount factor. The household budget constraint is

\[ (1 - \tau_{l,t}) w_t l_t + (1 - \tau_{k,t}) r_t k_t + \pi_t + \tau_t = c_t + x_t, \]  

(4)

where \( \tau_{l,t} \) and \( \tau_{k,t} \) are distortionary labor and capital income taxes while \( \tau_t \) is the lump-sum government transfers. Investment \( x_t \) is defined by the capital accumulation law:

\[ n a k_{t+1} = x_t + (1 - \delta) k_t. \]  

(5)

The government collects distortionary taxes from the household in order to finance government consumption while the remainder is transferred to the household in a lump-sum fashion. Therefore, the government budget constraint is

\[ g_t + \tau_t = \tau_{l,t} w_t l_t + \tau_{k,t} r_t k_t. \]  

(6)

Combining the government budget constraint (6) and the household budget constraint (4) making use of the definition of profits (1), we obtain the resource constraint

\[ y_t = c_t + x_t + g_t. \]  

(7)

\(^6\)We assume the maximum work week as \( 14 \times 7 = 98 \) and normalize hours worked per worker \( h_t \) as

\[ h_t = \frac{\text{average work week}}{98} \]

which is bounded between 0 and 1. Therefore, the detrended labor

\[ l_t = \frac{\text{average work week total employment}}{98} \frac{\text{total employment}}{\text{total population}} \]

is also bounded between 0 and 1.
Labor and investment wedges \( \{\omega_{lt}, \omega_{kt}\} \) are defined as:

\[
\omega_{lt} = 1 - \tau_{lt}, \\
\omega_{kt} = 1 - \tau_{kt}.
\]

Technically speaking, \( \omega_{lt} \) drives a wedge between the consumption-leisure marginal rate of substitution and the marginal product of labor while \( \omega_{kt} \) drives a wedge between the intertemporal marginal rate of substitution and the marginal return on investment. For convenience, we define government consumption wedges as the deviation of government purchases from its steady state level:

\[
\omega_{g,t} = g_t - g_t^g.
\]

\[8\]

### 2.3 Wedges

We define the efficiency, government consumption, investment and labor wedges \( \omega_t = (\omega_{e,t}, \omega_{g,t}, \omega_{k,t}, \omega_{l,t})' \) such that an increase in each wedge should lead to an increase in output. Increases in efficiency wedge directly increases production and stimulates factor demand by increasing the marginal product of inputs. On the other hand, increases in labor and investment wedges stimulate output by encouraging the household to increase supply of factor inputs through an increase in the marginal income associated with them. Therefore we refer to increases in efficiency, investment and labor wedges as “improvements”. High government consumption wedges should also increase output due to the increase in aggregate demand. However, we do not call an increase in government consumption as an “improvement” since this is associated with the crowding-out of household consumption and investment, which leads to household welfare deterioration. Following CKM (2007), we assume that the wedges are exogenous and follow a stochastic process. Defining a vector of log-linearized wedges, \( \tilde{\omega}_t = (\tilde{\omega}_{e,t}, \tilde{\omega}_{g,t}, \tilde{\omega}_{k,t}, \tilde{\omega}_{l,t})' \) where \( \tilde{\omega}_t = \ln \omega_t - \ln \omega \), we assume that the wedges follow a first order VAR process:

\[
\tilde{\omega}_t = P \tilde{\omega}_{t-1} + \epsilon_t
\]

\[9\]

where \( \epsilon_t = (\epsilon_{e,t}, \epsilon_{g,t}, \epsilon_{k,t}, \epsilon_{l,t})' \) are innovations to the wedges. Following CKM (2007) we allow spill-over of wedges through \( P \) and contemporaneous correlations of innovations in \( V \).

### 2.4 Equilibrium

The competitive equilibrium is given by a price vector \( \{r_t, w_t\} \) and an allocation of quantities \( \{y_t, c_t, x_t, l_t, z_t, g_t, \tau_t, \omega_{e,t}, \omega_{g,t}, \omega_{k,t}, \omega_{l,t}\} \) such that: (a) the household
maximizes utility given \( \{r_t, w_t, \tau_t, \omega_{k,t}, \omega_{l,t}\} \); (b) the firm maximizes profits given \( \{r_t, w_t, z_t\} \); (c) the government budget constraint (6) and the resource constraint (7) holds; and (d) the wedges follow the stochastic process (9). The competitive equilibrium is characterized by a set of first-order conditions given by: (a) the Euler equation (first order condition with respect to capital) equalizing present discounted value of marginal utility of future consumption to its marginal cost:

\[
\frac{1}{c_t} = \frac{\beta}{na} E_t \left[ \frac{1}{c_{t+1}} \left( \omega_{k,t+1} \theta \frac{y_{t+1}}{k_{t+1}} + 1 - \delta \right) \right],
\]

(b) the first-order equation with respect to labor equating marginal rate of substitution between consumption and leisure to the marginal product of labor:

\[
\frac{1 - \Psi}{\Psi} \frac{c_t}{1 - l_t} = \omega_{l,t} (1 - \theta) \frac{y_t}{l_t},
\]

(c) the resource constraint (7) given (8), (d) the capital law of motion (5), and (e) the production function (2) given (3).

3 Quantitative Analysis

3.1 Parameter Values

The first step in BCA implementation is to obtain the parameters of the model through usual calibration techniques for each country. For calibration purposes, we assume that there are no distortions in the steady state so that \( \omega = \{1, 1, 1, 1\} \). Capital share \( \theta \) is calibrated to match the capital income share derived from data. The productivity growth trend \( a \) is computed as the average growth rate of per capita output. Population growth trend \( n \) is directly computed from adult population data\(^7\).

We construct the total capital stock series as the sum of net fixed capital stock and household durables in order to compute the total annual depreciation rate \( \delta \). The subjective discount factor \( \beta \) is calibrated using the steady state capital Euler equation (10) to match steady state capital-output ratio given the productivity growth trend \( a \), population growth \( n \), capital share \( \theta \) and the depreciation rate \( \delta \). The preference weight \( \Psi \) is calibrated using the steady state labor first order condition (11) given the capital share \( \theta \), to match the steady state consumption-output ratio and the steady state labor. The values are listed in Table 3.

Once we have the calibrated parameters, the next step is to estimate the stochastic process of the wedges (9) for which we employ the Bayesian techniques. Structural estimation is necessary for the business cycle accounting procedure since investment

\(^7\)We used total population for China since we do not have adult population data.
wedges are defined in the intertemporal equilibrium condition (10) that depends on expectations about the future state of the economy which is not directly observable. The estimated parameters are the lag parameters in $P$, the standard deviation of the errors, and the cross-correlations between the errors in $V$. Since there are 4 exogenous variables, we use the time series data of output, consumption, investment and labor as observables. The Bayesian priors and the parameters of the vector and the point estimates of these parameters are listed in the appendix.

3.2 Simulation

The first step in the simulation process is to solve the model for linear decision rules for linearized endogenous variables $\tilde{k}_{t+1}$ and $\tilde{q}_t = (\tilde{y}_t, \tilde{c}_t, \tilde{x}_t, \tilde{b}_t)'$:

$$
\begin{align*}
\tilde{k}_{t+1} &= A\tilde{k}_t + B\tilde{\omega}_t, \\
\tilde{q}_t &= C\tilde{k}_t + D\tilde{\omega}_t.
\end{align*}
$$

Note that, given observed investment, the entire series of $\tilde{k}_t$ can be directly generated using the perpetual inventory method (assuming an initial value $\tilde{k}_0 = 0$):

$$
\tilde{k}_{t+1} = \frac{x}{nak} \tilde{x}_t + \frac{1 - \delta}{nak} \tilde{k}_t,
$$

Then the wedges can be computed as

$$
\tilde{\omega}_t = D^{-1}\left(\tilde{q}_t - C\tilde{k}_t\right).
$$

Once the wedges are computed, they are used for simulation in step 2. We compute the endogenous reaction of selected variables to the changes in a chosen wedge $\tilde{\omega}_{j,t}$ by plugging its time series into the linear decision rules of endogenous variables:

$$
\begin{align*}
\tilde{k}_{t+1}^{\tilde{\omega}_j} &= A\tilde{k}_t^{\tilde{\omega}_j} + B\tilde{\omega}_{j,t}^{\tilde{\omega}_j}, \\
\tilde{q}_t^{\tilde{\omega}_j} &= C\tilde{k}_t^{\tilde{\omega}_j} + D\tilde{\omega}_{j,t}^{\tilde{\omega}_j}.
\end{align*}
$$

By definition, plugging in all wedges into the model will exactly reproduce the observable data:

$$
\tilde{q}_t^{\tilde{\omega}} = C\tilde{k}_t + D\tilde{\omega}_t = C\tilde{k}_t + DD^{-1}\left(\tilde{q}_t - C\tilde{k}_t\right) = \tilde{q}_t.
$$

Therefore, we can easily decompose the effects of each wedges on the observables due to linearity of the decision rules:

$$
\tilde{q}_t^{\tilde{\omega}_x} + \tilde{q}_t^{\tilde{\omega}_y} + \tilde{q}_t^{\tilde{\omega}_k} + \tilde{q}_t^{\tilde{\omega}_i} = \tilde{q}_t^{\tilde{\omega}}.
$$
3.3 Results

Figure 1 presents the linearly detrended macroeconomic variables in Brazil, China, India and Russia for our sample period of 1990 – 2009\textsuperscript{8}. The detailed sources and data construction methods are listed in the data appendix. In reporting our results, we show the log deviations of the variables with respect to the steady state (where the first year of data availability is taken as the steady state).

Figure 2 plots the time paths of output and computed wedges for each country. For the most part, we do not find much commonality in wedge movements in the four nations. For example, while efficiency wedges have been above the trend in Brazil and Russia throughout the entire period, it has been below trend for most of the time in India and China. In Brazil, there was a temporary slow down in the growth of efficiency during 1997 – 2003. In Russia, it took off in 1998 and kept growing at an enormous rate, suggesting a positive impact of efficiency on growth. In India, while efficiency wedges temporarily improved in 2005, since then it has suddenly collapsed. In China, while efficiency wedges deteriorated during the 1995 – 2001 period, it shows a gradually improvement ever since. In contrast, in India, except for a small uptick during 2003 – 2005, efficiency has been below trend. It is hard to find common patterns in government consumption wedges and labor wedges as well, except for China and Brazil that saw an improvement in government consumption wedge during mid-twenties. Perhaps the common thread amongst all four nations is the evolution of investment wedges in the last decade. Investment wedges have been below the trend in Brazil and Russia and above trend in India and China throughout the entire period. However, they show improvements in all countries during the 2000s, a common factor in an otherwise diverse experience of the BRICs. This suggests that improvements in investment market frictions potentially aided the resurgence of BRICs since the mid-2000s.

In Table 4, we report the standard deviation of wedges with respect to output and the correlations of wedges with output for various leads and lags\textsuperscript{9} to ascertain

\textsuperscript{8}The variables are plotted as log deviations from their 1990 value (1992 in case of Russia).
\textsuperscript{9}As defined in CKM (2007), a "$k$ - th lag" is the correlation between the $t - k$ th value of the variable of interest with output at period $t$. 

10
various comovements. A positive correlation indicates a positive association between a given wedge and the observed economic outcome, and vice versa. Efficiency wedges, for the most part, are positively correlated with output in all countries except India, where the correlation turns negative contemporaneously and for the leads +1 and +2. Investment wedges also show a positive correlation with output in all countries. Labor wedges are positively correlated with output in Brazil and Russia, but negatively correlated in India. In China, while labor wedges become positively correlated for contemporaneous periods and leads +1, +2, the magnitude remains low. As for government consumption wedges, while they are positively correlated with output in Brazil (with the exception of the leads +1, +2), in India, and China, they are negatively correlated with output in Russia for all leads and lags. Given our wedges, we next feed them one by one in our benchmark model and simulate output. 

Table 5 presents the decomposition of the impact of each wedges on output and the investment to output ratio. We define a contribution indicator of each wedge $\omega_j$ on an endogenous variable $v$ as:

$$
cont_j = \text{corr}(\tilde{v}_t^{\omega_j}, \tilde{v}_t) * \frac{\text{std}(\tilde{v}_t^{\omega_j})}{\text{std}(\tilde{v}_t)}
= \frac{\text{cov}(\tilde{v}_t^{\omega_j}, \tilde{v}_t)}{\text{var}(\tilde{v}_t)}.
$$

Due to linearity,

$$
\sum_j cont_j = 1,
$$

as described in Otsu (2010b). Therefore, we can consider the value of the indicator as the contribution of each wedge to the fluctuation of the variable of interest.

### 3.3.1 Benchmark Model

First, we provide the simulation results for output in Table 5 (plot of simulated output in Figure 3). Since the economies grew particularly rapidly since 2000, we also specifically discuss the period 2000 to 2009.

In Brazil, efficiency, investment and labor wedges all contribute significantly explaining 29.3%, 36.8%, and 49.0% of output fluctuations respectively. Efficiency wedges are particularly significant in the 2000s with a contribution of 93.2%, while the contributions of investment and labor wedges, though positive, are much lower. As the figure depicts, the model with only efficiency wedges while capturing the short
run output fluctuation quite well, predicts a much higher output level throughout the entire period than witnessed in the data. By 2009, the model predicts output to be 13 percentage points above the trend. The growth in output that would have materialized with efficiency wedges alone are tempered by government consumption wedge. Investment and labor wedges for their part account for the sub-par economic performance of the 1990s and marginally contribute to the recovery of the 2000s. In Russia, during the overall sample period, efficiency wedges have a contribution higher than 100% while all other wedges have negative contributions. According to the figure, this is because the model with only efficiency wedges predicts the economy to recover much faster from the recession in the 1990s and grow much faster in the 2000s than it actually did. On the other hand, investment wedges predict a decline in output throughout the entire period. Therefore, investment wedges contribute to the downturn in 1990s while efficiency wedges aid Russia in recuperating much of the output loss in the 1990s to get back on the development track.

In India, investment wedges contribute the most to the fluctuation of output with an overall contribution of 87.4% over the entire period. This is mainly because of the 2000s where the contribution of investment wedge rises to 105.4%. Interestingly, during the 1990s the contribution of efficiency wedge at 79.6% was much higher than that of the investment wedge at 26.5%. When we run the model with only efficiency wedge, it performs quite well in predicting the fluctuation in output until 2005. However, it fails to predict the rapid growth after 2005. This is where the investment wedge comes in and investment wedges alone do a better job of accounting for the rapid acceleration of Indian growth during the 2000s well to the sample end. China presents a similar picture with efficiency wedges being the most important force in accounting for the output movement with a contribution of 72.6%. However, during the 2000s the contribution of investment wedges, 72.0%, becomes larger than that of efficiency wedges, 41.5%. According to the figure, the model with only efficiency wedges can almost perfectly reproduce the output fluctuations until 2004. However, mirroring the experience of India, it fails to account for the further rapid growth after 2004. On the other hand, investment wedges have significant impacts on output fluctuation throughout the entire 2000s till the end of the sample period, much like in India.

The unique experience of each country nevertheless show some common patterns, particularly in the last decade. While Brazilian and Russian growth was facilitated primarily by improvements in production efficiency (Brazil also benefitting to some extent from decline in investment market frictions), India and China grew primarily as a result of decline in investment market frictions, particularly in the later half of the 2000s, though, to some extent, China also benefitted from efficiency improvements as it did not experience the sudden loss of productive efficiency as India did since 2005. The contribution of labor and government consumption wedges to growth is
negligible in all four nations.

4 Sensitivity Analysis

4.1 Test 1: Efficiency Wedges as Productivity Growth

In CKM (2007) efficiency wedges are defined as temporary shocks to productivity. However, shocks to productivity might be permanent rather than temporary. Recall that in Figure 1, detrended output had fallen during the 1990s and then rapidly surged during the 2000s in all BRICs nations. In order to illustrate these medium term cycles better, it might be more appropriate to model efficiency wedges as shocks to the trend component of productivity rather than the cyclical component as suggested by Aguiar and Gopinath (2007). In this section, we alter the definition of efficiency wedges and compare the results to those in the benchmark model.

4.1.1 Model II

The only alteration we make from the benchmark model is the definition of efficiency wedges (3). First, we consider efficiency wedges as the growth in productivity between the previous period \((t - 1)\) and the current period \((t)\):

\[
\omega_{e,t} = \frac{\gamma_t}{\gamma_{t-1}}.
\]

We call this setting as model II. In model II, the realization of current productivity will define the growth of productivity and agents will anticipate the growth rate to gradually return to its mean according to (9) while this causes a permanent shift in the trend level. Therefore, the income effect caused by efficiency wedges should be stronger than that in the benchmark model\(^{10}\).

4.1.2 Model III

An alternative way to model efficiency wedges as productivity growth is to assume that current efficiency wedges lead to a growth in productivity between the current period \((t)\) and future period \((t + 1)\):

\[
\omega_{e,t} = \frac{\gamma_{t+1}}{\gamma_t}.
\]

\(^{10}\)In Aguiar and Gopinath (2007) there are shocks not only to the trend but also to the transitory component. The trend shock reflects the deviation of the productivity growth rate from its mean while the transitory component captures the deviation of the productivity from its trend level. Therefore, model II is equivalent to the Aguiar and Gopinath (2007) model without the transitory component.
We denote this setting as model III. In this model, the agents know the one-period-ahead productivity level when they make decisions on current choice variables. Also, as in model II, the agents will consider efficiency wedges as permanent shocks to the productivity level.

4.1.3 Simulation

Model II and Model III are estimated and simulated in a similar fashion as the prototype model. One important modification is that since we are defining efficiency wedges as shocks to the growth of productivity, we have to define the productivity level as an endogenous state variable. The linear decision rules of endogenous variables are:

\[
\begin{align*}
\widetilde{s}_{t+1} &= A\tilde{s}_t + B\widetilde{\omega}_t, \\
\tilde{q}_t &= C\tilde{s}_t + D\tilde{\omega}_t,
\end{align*}
\]

where we define the endogenous state variables \( \tilde{s}_t = (\tilde{k}_t, \tilde{A}_t) \). The entire series of \( \tilde{k}_t \) and \( \tilde{A}_t \) can be directly computed from

\[
\begin{align*}
\tilde{k}_{t+1} &= \frac{x}{nak} \tilde{x}_t + \frac{1 - \delta}{na} \tilde{k}_t, \\
\tilde{A}_t &= \frac{\tilde{y}_t}{1 - \theta} \frac{\theta \tilde{k}_t}{1 - \theta} - \tilde{l}_t,
\end{align*}
\]

assuming initial values \( \tilde{k}_0 = 0, \tilde{A}_0 = 0 \). Then the wedges can be computed as

\[
\widetilde{\omega}_t = D^{-1} (\tilde{q}_t - C\tilde{s}_t).
\]

Simulation is carried out in the same fashion as the benchmark model:

\[
\begin{align*}
\widetilde{s}_{t+1}^{\omega_j} &= A\tilde{s}_t^{\omega_j} + B\widetilde{\omega}_{j,t}, \\
\tilde{q}_t^{\omega_j} &= C\tilde{s}_t^{\omega_j} + D\tilde{\omega}_{j,t}.
\end{align*}
\]

4.1.4 Results

Since the growth shocks introduced in this section affects the expectations of the future, not only efficiency wedges but also investment wedges, that depend on expectations about future, are affected. The labor and government wedges are exactly the same as in the benchmark model. The output decomposition is plotted in Figure 4 and Table 6 provides the magnitudes.
The simulation results under the alternative models turn out to be similar to those in the benchmark model for the most part. In Brazil, under both the alternative specifications, investment and labor wedges account for the stagnation in the 1990s while efficiency wedges are important in accounting for the rapid growth in the 2000s. In Russia, investment wedges cause the downturn during the 1990s while efficiency wedges salvage the economy in the 2000s. In India, efficiency wedges account for the output fluctuations up to the mid-2000s while investment wedges are important in accounting for the rapid growth in the later 2000s. In China, efficiency wedges play a very important role in accounting for output fluctuations in both decades. The contribution of investment wedges during the 2000s for model II and III, 35.8% and 20.6% respectively, are considerably lower compared to that in the benchmark model, 72.0%. Government consumption wedges have higher contribution than in the benchmark model to compensate for this. Nonetheless, investment wedges still play an important role in the rapid growth during the later 2000s. It is important to note that the quantitative impact of the efficiency wedges are quite similar across the three models. Intuitively speaking, changing the definition of efficiency wedges does not change the realizations of productivity $A_t$ but it affects the expectations on future productivity. The result that the effects of efficiency wedges on output are robust across the three models indicates that the effects of the realization of productivity is more important than the expectations they generate.

4.2 Test 2: Benchmark Model with Investment Adjustment Costs

In the benchmark model capital stock is accumulated following the capital law of motion (5). However, as CKM (2007) argues, investment adjustment costs can reflect costs in converting output to capital in a detailed model, or financial frictions can manifest themselves as investment adjustment costs in a prototype RBC model. How does this modification affect our results? The only equation that changes is the capital accumulation equation:

$$nak_{t+1} = x_t + (1 - \delta)k_t - \Phi \left( \frac{x_t}{k_t} \right) k_t$$

where

$$\Phi \left( \frac{x_t}{k_t} \right) = \frac{\phi}{2} \left( \frac{x_t}{k_t} - \lambda \right)^2.$$

The constant $\lambda$ is set at $\lambda = na - (1 - \delta)$ so that the adjustment cost is equal to zero in the steady state. The parameter $\phi$ is calibrated to match the marginal Tobin’s $Q$. 

15
to one:
\[
\frac{d \log q}{d \log (x/k)} = 1,
\]
where \( q \) is the effective price of investment relative to consumption:
\[
q = \frac{1}{1 - \phi}.
\]
This leads to \( \phi = \frac{k}{x} \). We plot the simulations of output under each of the four wedges in Figure 5 (we also plot the results of the benchmark model for comparison). Output decompositions are presented in Table 7.

> Table 7 about here>
> Figure 5 about here>

While our basic results do not change with efficiency and investment wedges playing an important role in the output recovery since 2000, some subtle differences are noted, especially regarding the role of government consumption wedge. During the period 2000 to 2009, the contribution of government consumption wedge to output fluctuations increase as compared to the benchmark model in India and China. However, it is still smaller in magnitude as compared to investment wedge. A higher contribution of government wedge also implies a lower contribution of efficiency wedge in China, as compared to the benchmark model, though still coming in second to investment wedge in terms of its contribution.

5 Discussion: Decomposition, Wedges and Policies

The accounting results of the previous section highlight the importance of efficiency and investment wedges in output fluctuations. In this section, we take a look at some policy changes and institutional reforms that are consistent with the observed movement of these wedges. Our discussion mainly focuses on the 2000s due to data availability. Analytically, it works for us since it is the 2000s when we witness a sharp turnaround in growth of the BRIC nations.

Figure 6a plots the private credit share in GDP and the net FDI inflow to GDP ratio and suggests an increase in both till 2008 when FDI declined as a result of the global downturn. Interestingly, domestic credit to the private sector did not show any such decline. Increased capital flows suggest an improvement in credit worthiness borne out by the financial market indicators (Figure 6b) provided by the IMD World Competitiveness Yearbook (henceforth, WCY). There is an improvement in
credit rating, credit availability as well as the perception of businesses as to how encouraging the cost of capital was in the economy for all BRIC nations. These improvements are consistent with improved investment wedges that would lead to capital inflows fueled by rising credit ratings and increased the availability of capital for domestic businesses. Financial development is also consistent with observed production efficiency. On one hand, an increase in production efficiency should increase capital inflows as higher (perceived) efficiency leads to higher expected growth and lower probabilities of default, which is reflected in the rise in the country credit ratings. On the other hand, an increase in capital inflows can affect production efficiency through various channels. First, as discussed in Findlay (1978), an increase in FDI inflows could generate productivity spillovers through the import of managerial and organizational capital from foreign firms with superior efficiency. This effect could be particularly important in the banking sector as it improves the domestic resource allocation and thus the economy-wide efficiency. Next, as shown in Obstfeld (1994), greater diversification of income risk can lead to production specialization and the pursuit of riskier investment projects with high expected return. Finally, as discussed in Rajan and Zingales (2003), international financial integration will impose discipline on macroeconomic policies as transparency and good governance is essential to attract foreign capital and avoid capital flight. Financial liberalization and the resulting development in the financial market is consistent with the observed improvement in investment wedges in our model. When investment wedges are low, the expected return on investment is high relative to the intertemporal marginal rate of substitution as shown in (10). This can be caused by investment market distortions such as interest rate controls or capital controls which hampers the efficient flow of capital from the households to the firms. Financial liberalization increases the availability of capital by removing these distortions and enables firms to seize profitable investment opportunities. As a result, investment rises which brings down the expected return on investment due to diminishing marginal product of capital. Therefore, the gap between the intertemporal marginal rate of substitution and the expected return on capital should shrink.

\[\text{Figure 6a about here}\]

\[\text{Figure 6b about here}\]

Next, we track some institutional and governance indicators that provide the necessary framework for successful financial development and growth. Since our focus is to trace the development of BRIC policies over time, we focus on six time-series measures considered as conducive to economic development (definitions and expla-
nations are in the appendix). Figure 6c plots the six indices\(^{11}\) over time for each BRIC country and compare them to US standards where the measure ranges from \(-2.5\) (weak) to \(+2.5\) (strong). While it is clear that not all the indices show positive comovements with the time series of the estimated wedges, the two exceptions are government effectiveness and political stability to some extent. BRIC nations registered considerable improvement in government effectiveness particularly since early 2000s, though still below US standards. The indices in almost all instances move from negative to positive with almost doubling of the index value between 1996 and 2009. Even in case of Russia that scores the lowest, a 30% improvement in score is witnessed during the last decade. This translates to a 10 – rank climb in percentile ranks for all nations, with the exception of India that just climbs two spots. In terms of political stability, which is related to non violence and absence of terrorism, we witness a decline in 1990s till about mid-2000s when there is a turn-around. Brazil, the top scorer earns a score of \(-0.1\) (still in negatives though an improvement from \(-0.35\) in the 1990s). The most improvement was noticed in Russia that came out of the turbulent political transition of the 1990s to a more favorable domestic political climate. India is the only nation which seems to lag behind, not surprisingly due to its continued vulnerability to terrorism. Overall, we find that while some indices of improvement in institutional and political setup are consistent with our observed increases in productivity and investment wedges, not all indices reflect improvement.

<Figure 6c about here>

An interesting question would be why financial development might have impacted growth in efficiency in Brazil and Russia to a greater extent than in India and China, which particularly becomes apparent after 2004\(^{12}\). One important difference in these economies is the development stage that they were at when the reforms commenced. Brazil and Russia were coming out of a stagnation in early 2000s while India and China were already on the stable growth track since the 1990s\(^{13}\). Therefore, it might be the case that in Brazil and Russia, the impact of financial development on growth is much stronger - a case of catching up - as compared to India and China which were already on a stable development track\(^{14}\). India, in particular, is an aberration where

\(^{11}\)Voice & Accountability, Political Stability & Non Violence, Government Effectiveness, Regulatory Quality, Rule of Law, Control of Corruption

\(^{12}\)Bollard, Klenow and Sharma (2012) also find that FDI liberalization had little effect on the TFP growth in Indian manufacturing firms during the 1993 – 2007 period.

\(^{13}\)The growth trends in Brazil, Russia, India and China shown in Table 3 are 1.0%, 1.8%, 4.1% and 7.4% respectively.

\(^{14}\)Gente, Nourry and Leon-Ledesma (2012) show that financial liberalization can have positive or negative impacts on productivity growth depending on the national savings level in an endogenous growth setting with human capital accumulation.
efficiency suddenly collapsed after mid-2000s and we conjecture that the positive impact of financial development was overwhelmed by other factors that caused the efficiency collapse.

6 Conclusion

The growth of the BRIC nations - Brazil, Russia, India and China, has garnered much attention in the last decade. In this paper, we apply the Business Cycle Accounting methodology of Chari, Kehoe and McGrattan (2007) to explore the role of productivity fluctuations and changes in factor market distortions in accounting for the observed output fluctuations over the period 1990 to 2009. Our results, which are robust to methodological alternations, as well as model modifications, show that while each nations’ experience was unique, Brazil and Russia benefitted mostly from improved efficiency. India and China, on the other hand, saw a growth spurt in 2000s that can be largely accounted for by improvements in investment wedges, particularly in the latter half. Financial market developments in the BRIC economies, like increased credit flow aided by improved credit rating and business confidence are particularly consistent with improvements in efficiency and investment wedges. Indices denoting political stabilization and government effectiveness also improve possibly aiding efficiency gains and decline in investment market frictions.

One remaining question is why in Brazil and Russia financial development was accompanied by an improvement in efficiency while in India and China it was not. While we document that it relates to the development stage- Brazil and Russia coming out of a crisis to play catch-up and India and China already on a stable path- we leave further analysis of this topic for future research. According to institutional and governance indicators, BRIC nations have a long way to go before they catch up with the US standards. BRIC countries have taken steps in this direction by signing an accord to boost credit for trade transactions and authorizing establishment of a multilateral bank for funding projects in the developing world in the latest BRIC summit on March 29, 2012 with hopes of further such initiatives in the 2013 annual meeting of the BRICS.

References


21


Table 1: GDP ranking by PPP methodology (% share in world GDP)

*Source: International Monetary Fund Statistics*

<table>
<thead>
<tr>
<th>Year</th>
<th>1st</th>
<th>2nd</th>
<th>3rd</th>
<th>4th</th>
<th>5th</th>
<th>6th</th>
<th>7th</th>
<th>8th</th>
<th>9th</th>
<th>10th</th>
</tr>
</thead>
<tbody>
<tr>
<td>2011</td>
<td>U.S.</td>
<td>China</td>
<td>India</td>
<td>Japan</td>
<td>Germany</td>
<td>Russia</td>
<td>Brazil</td>
<td>U.K.</td>
<td>France</td>
<td>Italy</td>
</tr>
<tr>
<td>(19.11)</td>
<td>(14.36)</td>
<td>(5.67)</td>
<td>(5.58)</td>
<td>(3.92)</td>
<td>(3.02)</td>
<td>(2.93)</td>
<td>(2.86)</td>
<td>(2.81)</td>
<td>(2.32)</td>
<td></td>
</tr>
<tr>
<td>2010</td>
<td>U.S.</td>
<td>China</td>
<td>Japan</td>
<td>India</td>
<td>Germany</td>
<td>Russia</td>
<td>U.K.</td>
<td>Brazil</td>
<td>France</td>
<td>Italy</td>
</tr>
<tr>
<td>(19.53)</td>
<td>(13.61)</td>
<td>(5.81)</td>
<td>(5.46)</td>
<td>(3.96)</td>
<td>(3.00)</td>
<td>(2.93)</td>
<td>(2.93)</td>
<td>(2.87)</td>
<td>(2.39)</td>
<td></td>
</tr>
<tr>
<td>2005</td>
<td>U.S.</td>
<td>China</td>
<td>Japan</td>
<td>Germany</td>
<td>India</td>
<td>U.K.</td>
<td>France</td>
<td>Russia</td>
<td>Italy</td>
<td>Brazil</td>
</tr>
<tr>
<td>(22.26)</td>
<td>(9.46)</td>
<td>(6.83)</td>
<td>(4.40)</td>
<td>(4.29)</td>
<td>(3.41)</td>
<td>(3.28)</td>
<td>(2.99)</td>
<td>(2.88)</td>
<td>(2.80)</td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td>U.S.</td>
<td>Japan</td>
<td>China</td>
<td>Germany</td>
<td>India</td>
<td>France</td>
<td>U.K.</td>
<td>Italy</td>
<td>Brazil</td>
<td>Russia</td>
</tr>
<tr>
<td>(23.55)</td>
<td>(7.61)</td>
<td>(7.14)</td>
<td>(5.07)</td>
<td>(3.72)</td>
<td>(3.63)</td>
<td>(3.59)</td>
<td>(3.31)</td>
<td>(2.92)</td>
<td>(2.65)</td>
<td></td>
</tr>
<tr>
<td>1995</td>
<td>U.S.</td>
<td>Japan</td>
<td>China</td>
<td>Germany</td>
<td>France</td>
<td>U.K.</td>
<td>Italy</td>
<td>India</td>
<td>Brazil</td>
<td>Russia</td>
</tr>
<tr>
<td>(22.89)</td>
<td>(8.71)</td>
<td>(5.67)</td>
<td>(5.55)</td>
<td>(3.81)</td>
<td>(3.64)</td>
<td>(3.61)</td>
<td>(3.31)</td>
<td>(3.17)</td>
<td>(2.94)</td>
<td></td>
</tr>
<tr>
<td>1990</td>
<td>U.S.</td>
<td>Japan</td>
<td>Germany</td>
<td>France</td>
<td>Italy</td>
<td>U.K.</td>
<td>China</td>
<td>Brazil</td>
<td>India</td>
<td>Mexico</td>
</tr>
<tr>
<td>1985</td>
<td>U.S.</td>
<td>Japan</td>
<td>Germany</td>
<td>France</td>
<td>Italy</td>
<td>U.K.</td>
<td>Brazil</td>
<td>China</td>
<td>Mexico</td>
<td>India</td>
</tr>
<tr>
<td>(25.19)</td>
<td>(9.29)</td>
<td>(6.22)</td>
<td>(4.47)</td>
<td>(4.25)</td>
<td>(4.16)</td>
<td>(3.61)</td>
<td>(3.18)</td>
<td>(2.85)</td>
<td>(2.84)</td>
<td></td>
</tr>
<tr>
<td>1980</td>
<td>U.S.</td>
<td>Japan</td>
<td>Germany</td>
<td>France</td>
<td>Italy</td>
<td>U.K.</td>
<td>Brazil</td>
<td>Mexico</td>
<td>India</td>
<td>Spain</td>
</tr>
<tr>
<td>(24.64)</td>
<td>(8.65)</td>
<td>(6.74)</td>
<td>(4.74)</td>
<td>(4.48)</td>
<td>(4.28)</td>
<td>(3.92)</td>
<td>(2.97)</td>
<td>(2.53)</td>
<td>(2.41)</td>
<td></td>
</tr>
</tbody>
</table>
Table 2: Aggregate GDP and GDP per capita growth rates

*Data Source: World Bank and Penn World Tables*

Column (1) summarizes growth in Aggregate GDP while column (2) summarizes growth in GDP per capita.

<table>
<thead>
<tr>
<th></th>
<th>1960s</th>
<th>1970s</th>
<th>1980s</th>
<th>1990s</th>
<th>2000s</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(1)</td>
<td>(2)</td>
<td>(1)</td>
</tr>
<tr>
<td>U.S.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>4.66%</td>
<td>3.33%</td>
<td>3.32%</td>
<td>2.24%</td>
<td>3.04%</td>
</tr>
<tr>
<td>St. Dev.</td>
<td>(1.68%)</td>
<td>(1.67%)</td>
<td>(2.58%)</td>
<td>2.56%</td>
<td>(2.55%)</td>
</tr>
<tr>
<td>OECD</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>5.74%</td>
<td>4.42%</td>
<td>3.73%</td>
<td>2.67%</td>
<td>2.94%</td>
</tr>
<tr>
<td>St. Dev.</td>
<td>(0.74%)</td>
<td>(0.81%)</td>
<td>(1.89%)</td>
<td>(1.91%)</td>
<td>(1.44%)</td>
</tr>
<tr>
<td>Brazil</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>5.90%</td>
<td>2.97%</td>
<td>8.47%</td>
<td>5.92%</td>
<td>2.99%</td>
</tr>
<tr>
<td>St. Dev.</td>
<td>(3.68%)</td>
<td>(3.68%)</td>
<td>(3.48%)</td>
<td>(3.39%)</td>
<td>(4.76%)</td>
</tr>
<tr>
<td>Russia</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>St. Dev.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>India</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>6.67%</td>
<td>4.44%</td>
<td>2.93%</td>
<td>0.55%</td>
<td>5.69%</td>
</tr>
<tr>
<td>St. Dev.</td>
<td>(6.14%)</td>
<td>(6.01%)</td>
<td>(4.16%)</td>
<td>(4.06%)</td>
<td>(1.88%)</td>
</tr>
<tr>
<td>China</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>3.02%</td>
<td>0.89%</td>
<td>7.44%</td>
<td>5.34%</td>
<td>9.75%</td>
</tr>
<tr>
<td>St. Dev.</td>
<td>(14.85%)</td>
<td>(13.74%)</td>
<td>(5.62%)</td>
<td>(5.37%)</td>
<td>(3.24%)</td>
</tr>
</tbody>
</table>
Table 3. Parameters and Steady States  
*Source: Authors’ calculations*

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Explanation</th>
<th>Brazil</th>
<th>Russia</th>
<th>India</th>
<th>China</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a$</td>
<td>Average growth rate of per capita output</td>
<td>1.010</td>
<td>1.018</td>
<td>1.041</td>
<td>1.074</td>
</tr>
<tr>
<td>$n$</td>
<td>Average growth rate of population</td>
<td>1.017</td>
<td>0.999</td>
<td>1.019</td>
<td>1.007</td>
</tr>
<tr>
<td>$\theta$</td>
<td>Share of capital in output</td>
<td>0.521</td>
<td>0.526</td>
<td>0.713</td>
<td>0.293</td>
</tr>
<tr>
<td>$\delta$</td>
<td>Rate of depreciation</td>
<td>0.120</td>
<td>0.094</td>
<td>0.121</td>
<td>0.117</td>
</tr>
<tr>
<td>$\beta$</td>
<td>Subjective discount factor</td>
<td>0.849</td>
<td>0.939</td>
<td>0.776</td>
<td>1.042</td>
</tr>
<tr>
<td>$\Psi$</td>
<td>Elasticity of substitution between consumption and leisure</td>
<td>0.273</td>
<td>0.177</td>
<td>0.381</td>
<td>0.154</td>
</tr>
<tr>
<td>$y/k$</td>
<td>Steady state output to capital ratio</td>
<td>0.633</td>
<td>0.338</td>
<td>0.683</td>
<td>0.526</td>
</tr>
<tr>
<td>$l$</td>
<td>Steady state labor</td>
<td>0.230</td>
<td>0.193</td>
<td>0.218</td>
<td>0.230</td>
</tr>
<tr>
<td>$c/y$</td>
<td>Consumption as a share of output in the steady state</td>
<td>0.604</td>
<td>0.426</td>
<td>0.634</td>
<td>0.432</td>
</tr>
<tr>
<td>$x/y$</td>
<td>Investment as a share of output in the steady state</td>
<td>0.218</td>
<td>0.424</td>
<td>0.292</td>
<td>0.417</td>
</tr>
<tr>
<td>$g/y$</td>
<td>Government expenditure as a share of output in the steady state</td>
<td>0.179</td>
<td>0.150</td>
<td>0.074</td>
<td>0.151</td>
</tr>
</tbody>
</table>

**Benchmark model with Investment Adjustment Costs**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Explanation</th>
<th>Brazil</th>
<th>Russia</th>
<th>India</th>
<th>China</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\phi$</td>
<td>Sensitivity of investment to marginal $Q$</td>
<td>7.252</td>
<td>6.965</td>
<td>5.015</td>
<td>4.558</td>
</tr>
<tr>
<td>$\kappa$</td>
<td>Steady state investment to capital ratio</td>
<td>0.147</td>
<td>0.111</td>
<td>0.181</td>
<td>0.198</td>
</tr>
</tbody>
</table>
Table 4: Properties of the wedges

Source: Authors’ calculations

Benchmark Model

<table>
<thead>
<tr>
<th></th>
<th>Standard Deviation with respect to output</th>
<th>Cross Correlations of wedges with output at lag k=</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>−2</td>
</tr>
<tr>
<td><strong>BRAZIL</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Efficiency Wedges</td>
<td>2.43</td>
<td>0.24</td>
</tr>
<tr>
<td>Government Consumption Wedges</td>
<td>2.99</td>
<td>0.72</td>
</tr>
<tr>
<td>Investment Wedges</td>
<td>1.36</td>
<td>0.25</td>
</tr>
<tr>
<td>Labor Wedges</td>
<td>1.55</td>
<td>0.16</td>
</tr>
<tr>
<td><strong>RUSSIA</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Efficiency Wedges</td>
<td>7.61</td>
<td>0.87</td>
</tr>
<tr>
<td>Government Consumption Wedges</td>
<td>3.50</td>
<td>−0.28</td>
</tr>
<tr>
<td>Investment Wedges</td>
<td>9.61</td>
<td>−0.12</td>
</tr>
<tr>
<td>Labor Wedges</td>
<td>0.61</td>
<td>0.63</td>
</tr>
<tr>
<td><strong>INDIA</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Efficiency Wedges</td>
<td>2.16</td>
<td>0.43</td>
</tr>
<tr>
<td>Government Consumption Wedges</td>
<td>3.22</td>
<td>0.21</td>
</tr>
<tr>
<td>Investment Wedges</td>
<td>1.87</td>
<td>0.86</td>
</tr>
<tr>
<td>Labor Wedges</td>
<td>0.85</td>
<td>−0.55</td>
</tr>
<tr>
<td><strong>CHINA</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Efficiency Wedges</td>
<td>1.34</td>
<td>0.53</td>
</tr>
<tr>
<td>Government Consumption Wedges</td>
<td>3.55</td>
<td>0.54</td>
</tr>
<tr>
<td>Investment Wedges</td>
<td>1.64</td>
<td>0.24</td>
</tr>
<tr>
<td>Labor Wedges</td>
<td>1.48</td>
<td>−0.04</td>
</tr>
</tbody>
</table>
Table 5: Decomposition of Output - Benchmark Model

Source: Authors’ calculations

<table>
<thead>
<tr>
<th></th>
<th>Brazil</th>
<th>Russia</th>
<th>India</th>
<th>China</th>
</tr>
</thead>
<tbody>
<tr>
<td>1990:2009</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Efficiency Wedges</td>
<td>0.293</td>
<td>1.826</td>
<td>0.039</td>
<td>0.726</td>
</tr>
<tr>
<td>Government Consumption Wedges</td>
<td>−0.151</td>
<td>−0.196</td>
<td>0.014</td>
<td>0.049</td>
</tr>
<tr>
<td>Investment Wedges</td>
<td>0.368</td>
<td>−0.570</td>
<td>0.874</td>
<td>0.218</td>
</tr>
<tr>
<td>Labor Wedges</td>
<td>0.490</td>
<td>−0.060</td>
<td>0.073</td>
<td>0.006</td>
</tr>
<tr>
<td>1990:1999</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Efficiency Wedges</td>
<td>−0.535</td>
<td>−0.746</td>
<td>0.796</td>
<td>0.991</td>
</tr>
<tr>
<td>Government Consumption Wedges</td>
<td>−0.047</td>
<td>0.037</td>
<td>−0.118</td>
<td>−0.005</td>
</tr>
<tr>
<td>Investment Wedges</td>
<td>0.609</td>
<td>1.619</td>
<td>0.265</td>
<td>−0.142</td>
</tr>
<tr>
<td>Labor Wedges</td>
<td>0.973</td>
<td>0.090</td>
<td>0.057</td>
<td>0.155</td>
</tr>
<tr>
<td>2000:2009</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Efficiency Wedges</td>
<td>0.932</td>
<td>1.559</td>
<td>−0.128</td>
<td>0.415</td>
</tr>
<tr>
<td>Government Consumption Wedges</td>
<td>−0.153</td>
<td>−0.041</td>
<td>0.005</td>
<td>0.131</td>
</tr>
<tr>
<td>Investment Wedges</td>
<td>0.143</td>
<td>−0.437</td>
<td>1.054</td>
<td>0.720</td>
</tr>
<tr>
<td>Labor Wedges</td>
<td>0.078</td>
<td>−0.082</td>
<td>0.068</td>
<td>−0.266</td>
</tr>
</tbody>
</table>
Table 6: Decomposition of Output -Alternative Models

Source: Authors’ calculations

<table>
<thead>
<tr>
<th>Source: Authors’ calculations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Efficiency Wedges</td>
</tr>
<tr>
<td>0.239</td>
</tr>
<tr>
<td>Government Consumption Wedges</td>
</tr>
<tr>
<td>Investment Wedges</td>
</tr>
<tr>
<td>Labor Wedges</td>
</tr>
<tr>
<td>Efficiency Wedges</td>
</tr>
<tr>
<td>Government Consumption Wedges</td>
</tr>
<tr>
<td>Investment Wedges</td>
</tr>
<tr>
<td>Labor Wedges</td>
</tr>
<tr>
<td>Efficiency Wedges</td>
</tr>
<tr>
<td>Government Consumption Wedges</td>
</tr>
<tr>
<td>Investment Wedges</td>
</tr>
<tr>
<td>Labor Wedges</td>
</tr>
</tbody>
</table>
Table 7: Decomposition of Output - Benchmark Model with Investment Adjustment Costs

*Source: Authors’ calculations*

<table>
<thead>
<tr>
<th></th>
<th>Brazil</th>
<th>Russia</th>
<th>India</th>
<th>China</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1990:2009</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Efficiency Wedges</td>
<td>0.273</td>
<td>2.322</td>
<td>−0.166</td>
<td>0.636</td>
</tr>
<tr>
<td>Government Consumption Wedges</td>
<td>−0.052</td>
<td>−0.367</td>
<td>0.214</td>
<td>0.075</td>
</tr>
<tr>
<td>Investment Wedges</td>
<td>0.399</td>
<td>−0.941</td>
<td>0.579</td>
<td>0.288</td>
</tr>
<tr>
<td>Labor Wedges</td>
<td>0.380</td>
<td>−0.014</td>
<td>0.374</td>
<td>0.001</td>
</tr>
<tr>
<td><strong>1990:1999</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Efficiency Wedges</td>
<td>−0.651</td>
<td>−0.746</td>
<td>0.723</td>
<td>0.893</td>
</tr>
<tr>
<td>Government Consumption Wedges</td>
<td>0.030</td>
<td>−0.082</td>
<td>−0.283</td>
<td>−0.013</td>
</tr>
<tr>
<td>Investment Wedges</td>
<td>0.749</td>
<td>1.807</td>
<td>0.396</td>
<td>0.067</td>
</tr>
<tr>
<td>Labor Wedges</td>
<td>0.871</td>
<td>0.020</td>
<td>0.165</td>
<td>0.052</td>
</tr>
<tr>
<td><strong>2000:2009</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Efficiency Wedges</td>
<td>1.123</td>
<td>2.234</td>
<td>−0.331</td>
<td>0.298</td>
</tr>
<tr>
<td>Government Consumption Wedges</td>
<td>−0.015</td>
<td>−0.295</td>
<td>0.232</td>
<td>0.209</td>
</tr>
<tr>
<td>Investment Wedges</td>
<td>−0.002</td>
<td>−0.958</td>
<td>0.715</td>
<td>0.590</td>
</tr>
<tr>
<td>Labor Wedges</td>
<td>−0.106</td>
<td>0.018</td>
<td>0.384</td>
<td>−0.096</td>
</tr>
</tbody>
</table>
Figure 1: Real Macro Aggregates

Output

Investment

Consumption

Labor

Brazil Russia India China
Note: “Output (Y)” includes GDP and the imputed service flow from consumer durables. It is decomposed into "Consumption (C)" that consists of household consumption of non-durables and services (where the imputed service flow from consumer durables are included) and "Investment (X)" that includes gross domestic capital formation and household expenditures on consumer durables while the residual is defined as "Government Consumption (G)" so that Y=C+X+G "Labor (L)" represents total hours worked which consists of total employment and hours worked per workers. All variables are divided by the adult population. Output, consumption and investment are linearly detrended by the average per adult output growth rate over the 1990-2009 period setting 1990 at the trend level

Source: The data is primarily collected from the Penn World Tables edition 7.0 and its extension made by Duncan Foley
Figure 2: Estimated Wedges in the benchmark model

Note: Efficiency wedges in our benchmark model are estimated as shocks to the “level” of productivity.
Figure 3: Simulated Output in the benchmark model

- Brazil
- Russia
- India
- China

Legend:
- Efficiency
- Government
- Investment
- Labor
- Output
Figure 4: Simulated output under model II

Note: In model II, efficiency wedges are modeled as shocks to growth rate of realized productivity.
Figure 4 contd.: Simulated output under model III

Note: In model III, efficiency wedges are modeled as shocks to future productivity growth
Figure 5: Simulated output under benchmark model with investment adjustment costs

Output with efficiency wedges

Brazil

Russia

Output with government consumption wedges

India

China

Brazil

Russia

India

China
Figure 5 contd.: Simulated output under benchmark model with investment adjustment costs

Output with investment wedges

Output with labor wedges

Brazil

Russia

India

China

Brazil

Russia

India

China
Note: AC denotes the benchmark model with quadratic adjustment costs for investment, while the benchmark model is exactly similar to the AC model except without the quadratic adjustment costs. We feed in efficiency, government consumption, investment and labor wedges one at a time and compare the model simulations of output under the AC and benchmark model with that in the data.
Figure 6a: Flow of Domestic Credit to Private Sector and Inflows of FDI

Domestic Credit to the Private Sector (% of GDP)

Net FDI Inflows (% of GDP)
Figure 6b:  Financial Market Indicators

Country Credit Rating

Credit Availability for Businesses

Capital Affordability
Figure 6c: Measures of Institutional and Policy Reforms

Voice and Accountability

Political Stability

Government Effectiveness

Regulatory Quality

Rule of Law

Control of Corruption
1 Linearization Appendix

In this section we define the log-linearized equations of our model.

We define the log linearization of each detrended variables from their steady states as

\[ \tilde{v}_t = \ln \tilde{v}_t - \ln \bar{v} \]

Then the linearized equilibrium conditions are

\[ 0 = \frac{\beta}{na} k \tilde{k}_{t+1} - \frac{\beta}{na} \theta y \tilde{y}_{t+1} + \tilde{c}_{t+1} - \tilde{c}_t - \frac{\beta}{na} \theta y \tilde{g}_{k,t+1} \]

\[ 0 = \tilde{y}_t - \tilde{c}_t - \frac{1}{1 - l} \tilde{l}_t + \tilde{\omega}_{l,t} \]

\[ 0 = \tilde{y}_t - \frac{c}{y} \tilde{c}_t - \frac{x}{y} \tilde{x}_t - \frac{g}{y} \tilde{\omega}_{g,t} \]

\[ 0 = \bar{n} k \tilde{k}_{t+1} - \frac{x}{k} \tilde{x}_t - (1 - \delta) \tilde{k}_t \]

\[ 0 = \tilde{y}_t - \theta \tilde{k}_t - (1 - \theta) \tilde{g}_{t} - (1 - \theta) \tilde{l}_t \]
Finally, we consider three cases regarding the definition of $\hat{\omega}_{e,t}$. The first case follows Chari, Kehoe and McGrattan (2007) where efficiency wedges $\omega_{e,t}$ directly affect the level of productivity:

$$\hat{\omega}_{e,t} = \gamma_t.$$  
(Model I)

In the second case, we define efficiency wedges as the growth of productivity between the previous period and the current period:

$$\hat{\omega}_{e,t} = \gamma_t - \gamma_{t-1}.$$  
(Model II)

Finally, in the third case, we define efficiency wedges as the growth of productivity between the current period and the next period:

$$\hat{\omega}_{e,t} = \gamma_{t-1} - \gamma_t.$$  
(Model III)
2 Parameters of the Vector AR (1) Stochastic Process of the Wedges

Given the underlying vector AR(1) stochastic process for the wedges and the data on output, consumption, investment and labor in Brazil, Russia, India and China, we estimate the wedges using Bayesian techniques. The bayesian priors are listed in Table A. The parameters underlying the vector AR(1) process for the wedges in Brazil, Russia, India and China are listed in Table B for the benchmark model where productivity wedge is modeled as shocks to the level of productivity. Tables C and D list the parameters of the AR(1) process governing the shocks under models II and III where productivity wedges are modeled as shocks to the realized growth rate and future growth rate of productivity respectively.

Table A: The Bayesian Priors for structural estimation of wedges

<table>
<thead>
<tr>
<th></th>
<th>Prior Distribution</th>
<th>Prior Mean</th>
<th>Prior Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P$ Diagonal</td>
<td>Beta</td>
<td>0.8</td>
<td>0.2</td>
</tr>
<tr>
<td>$P$ Off-Diagonal</td>
<td>Normal</td>
<td>0</td>
<td>0.2</td>
</tr>
<tr>
<td>V Standard Deviation</td>
<td>Inverse Gamma</td>
<td>0.05</td>
<td>inf</td>
</tr>
<tr>
<td>V Correlation</td>
<td>Uniform</td>
<td>0</td>
<td>$-1,1$</td>
</tr>
</tbody>
</table>

Table B: Parameters of the Vector AR(1) Stochastic Process driving the wedges -Benchmark Model
<table>
<thead>
<tr>
<th>P</th>
<th>V</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brazil</td>
<td></td>
</tr>
<tr>
<td>0.7930</td>
<td>0.1990</td>
</tr>
<tr>
<td>−0.3500</td>
<td>0.7940</td>
</tr>
<tr>
<td>−0.0790</td>
<td>0.0200</td>
</tr>
<tr>
<td>−0.0070</td>
<td>−0.0510</td>
</tr>
<tr>
<td>Russia</td>
<td></td>
</tr>
<tr>
<td>0.9330</td>
<td>0.1890</td>
</tr>
<tr>
<td>−0.3470</td>
<td>0.8690</td>
</tr>
<tr>
<td>0.0390</td>
<td>−0.0410</td>
</tr>
<tr>
<td>0.0220</td>
<td>−0.0470</td>
</tr>
<tr>
<td>India</td>
<td></td>
</tr>
<tr>
<td>0.8440</td>
<td>0.0110</td>
</tr>
<tr>
<td>0.2390</td>
<td>0.7790</td>
</tr>
<tr>
<td>−0.0050</td>
<td>0.0050</td>
</tr>
<tr>
<td>−0.0080</td>
<td>0.0610</td>
</tr>
<tr>
<td>China</td>
<td></td>
</tr>
<tr>
<td>0.8250</td>
<td>0.0280</td>
</tr>
<tr>
<td>−0.0150</td>
<td>0.8690</td>
</tr>
<tr>
<td>−0.0110</td>
<td>0.0050</td>
</tr>
<tr>
<td>0.1070</td>
<td>0.0330</td>
</tr>
</tbody>
</table>
Table C: Parameters of the Vector AR(1) Stochastic Process driving the wedges -Model II

<table>
<thead>
<tr>
<th></th>
<th>P</th>
<th>V</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brazil</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.5490</td>
<td>0.0047</td>
</tr>
<tr>
<td></td>
<td>0.0266</td>
<td>0.8200</td>
</tr>
<tr>
<td></td>
<td>0.1770</td>
<td>0.0167</td>
</tr>
<tr>
<td></td>
<td>-0.0975</td>
<td>0.0753</td>
</tr>
<tr>
<td>Russia</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.5668</td>
<td>0.0464</td>
</tr>
<tr>
<td></td>
<td>-0.3264</td>
<td>0.7534</td>
</tr>
<tr>
<td></td>
<td>0.0796</td>
<td>-0.3596</td>
</tr>
<tr>
<td></td>
<td>-0.6895</td>
<td>0.2033</td>
</tr>
<tr>
<td>India</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.5906</td>
<td>-0.0294</td>
</tr>
<tr>
<td></td>
<td>0.2724</td>
<td>0.8427</td>
</tr>
<tr>
<td></td>
<td>-0.0013</td>
<td>-0.0005</td>
</tr>
<tr>
<td></td>
<td>-0.1240</td>
<td>0.0566</td>
</tr>
<tr>
<td>China</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.4931</td>
<td>-0.0009</td>
</tr>
<tr>
<td></td>
<td>0.6399</td>
<td>0.8373</td>
</tr>
<tr>
<td></td>
<td>0.7828</td>
<td>0.0157</td>
</tr>
<tr>
<td></td>
<td>-0.2634</td>
<td>0.0195</td>
</tr>
</tbody>
</table>
Table D: Parameters of the Vector AR(1) Stochastic Process driving the wedges -Model III

<table>
<thead>
<tr>
<th></th>
<th>Brazil</th>
<th>V</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>P</td>
<td>0.6078</td>
<td>0.1327</td>
<td>-0.4426</td>
<td>0.0945</td>
<td>0.0009</td>
<td>0.0000</td>
<td>0.0002</td>
</tr>
<tr>
<td></td>
<td>0.1489</td>
<td>0.7636</td>
<td>0.3575</td>
<td>-0.0446</td>
<td>0.0000</td>
<td>0.0019</td>
<td>0.0000</td>
</tr>
<tr>
<td></td>
<td>0.2001</td>
<td>-0.0175</td>
<td>0.7729</td>
<td>0.0463</td>
<td>0.0002</td>
<td>0.0000</td>
<td>0.0006</td>
</tr>
<tr>
<td></td>
<td>-0.3876</td>
<td>0.0148</td>
<td>0.5559</td>
<td>0.8092</td>
<td>-0.0001</td>
<td>-0.0002</td>
<td>0.0000</td>
</tr>
<tr>
<td>Russia</td>
<td>0.7895</td>
<td>0.0180</td>
<td>0.0654</td>
<td>0.2833</td>
<td>0.0085</td>
<td>0.0195</td>
<td>-0.0001</td>
</tr>
<tr>
<td></td>
<td>-0.1659</td>
<td>0.8373</td>
<td>-0.2054</td>
<td>0.2147</td>
<td>0.0195</td>
<td>0.0984</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>0.3150</td>
<td>-0.0646</td>
<td>0.9177</td>
<td>-0.1507</td>
<td>-0.0001</td>
<td>0.0000</td>
<td>0.0002</td>
</tr>
<tr>
<td></td>
<td>-0.2272</td>
<td>-0.0203</td>
<td>-0.1627</td>
<td>0.8610</td>
<td>-0.0010</td>
<td>-0.0041</td>
<td>-0.0001</td>
</tr>
<tr>
<td>India</td>
<td>0.5858</td>
<td>0.0152</td>
<td>-0.2131</td>
<td>0.2100</td>
<td>0.0150</td>
<td>0.000</td>
<td>0.0003</td>
</tr>
<tr>
<td></td>
<td>-0.0100</td>
<td>0.8537</td>
<td>0.1355</td>
<td>-0.1041</td>
<td>0.0000</td>
<td>0.0272</td>
<td>-0.0009</td>
</tr>
<tr>
<td></td>
<td>0.0170</td>
<td>-0.0003</td>
<td>0.9489</td>
<td>-0.2441</td>
<td>0.0003</td>
<td>-0.0009</td>
<td>0.0001</td>
</tr>
<tr>
<td></td>
<td>-0.1805</td>
<td>0.0476</td>
<td>-0.0699</td>
<td>0.8130</td>
<td>0.0007</td>
<td>-0.0007</td>
<td>-0.0001</td>
</tr>
<tr>
<td>China</td>
<td>0.7457</td>
<td>0.0259</td>
<td>-0.0299</td>
<td>0.0999</td>
<td>0.0020</td>
<td>0.0004</td>
<td>0.0084</td>
</tr>
<tr>
<td></td>
<td>1.1285</td>
<td>0.8002</td>
<td>-0.1127</td>
<td>-0.1722</td>
<td>0.0004</td>
<td>0.0110</td>
<td>-0.0022</td>
</tr>
<tr>
<td></td>
<td>-0.7322</td>
<td>0.1422</td>
<td>0.8069</td>
<td>0.2940</td>
<td>0.0084</td>
<td>-0.0022</td>
<td>0.0370</td>
</tr>
<tr>
<td></td>
<td>0.3301</td>
<td>-0.0192</td>
<td>-0.1401</td>
<td>0.8972</td>
<td>-0.0002</td>
<td>0.0003</td>
<td>-0.0011</td>
</tr>
</tbody>
</table>
3 Data Appendix

3.1 Data Sources

"Output (Y)" includes GDP and the imputed service flow from consumer durables. It is decomposed into "Consumption (C)" that consists of household consumption of non-durables and services (where the imputed service flow from consumer durables are included) and "Investment (X)" that includes gross domestic capital formation and household expenditures on consumer durables while the residual is defined as "Government Consumption (G)" so that \( Y = C + X + G \). "Labor (L)" represents total hours worked which consists of total employment and hours worked per workers. All variables are divided by the adult population\(^2\). Output, consumption and investment are linearly detrended by the average per adult output growth rate over the 1990 – 2009 period setting 1990 at the trend level\(^3\). The data is primarily collected from the Penn World Tables edition 7.0 and its extension made by Duncan Foley\(^4\). Table A1 presents the original sources of the data. PWT stands for Penn World Tables edition 7.1 and the extensions made by Duncan Foley. EM stands for the Eurominotor Global Market Information Database. ILO stands for the International Labor Organization LABORSTA database. The details of data construction follows.

<table>
<thead>
<tr>
<th>Table A1. Original Sources of the Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>GDP</td>
</tr>
<tr>
<td>Consumption share</td>
</tr>
<tr>
<td>Investment share</td>
</tr>
<tr>
<td>Employment</td>
</tr>
<tr>
<td>Hours worked per worker</td>
</tr>
<tr>
<td>Population</td>
</tr>
<tr>
<td>Adult Share in Total Population</td>
</tr>
<tr>
<td>Household Expenditure on Durables</td>
</tr>
<tr>
<td>Net fixed Capital Stock</td>
</tr>
<tr>
<td>Depreciation</td>
</tr>
<tr>
<td>Household Income Share of Capital</td>
</tr>
</tbody>
</table>

\(^1\)Therefore, \( G \) includes government purchases of goods and services as well as net exports. The inclusion of net exports in government consumption follows the tradition of a closed economy BCA model (Chari, Kehoe and McGrattan (2007)).

\(^2\)We use total population for China due to data availability.

\(^3\)Therefore, the output series will start at the trend level in 1990 and end at the trend level in 2009.

\(^4\)Source: https://sites.google.com/a/newschool.edu/duncan-foley-homepage/home/EPWT

\(^5\)For Russian capital stock and depreciation we refer to Izyumov and Vahaly (2008) because the Foley database reports capital stock data only for the 2004-2008 period.

\(^6\)Izyumov and Vahaly (2008) assume a constant 5% annual depreciation.
Employment $E$ is computed from the PWT data of GDP per capita ($rgdpl2$) and GDP per person counted in total employment ($rgdpl2te$) and population ($POP$):

$$E = \frac{rgdpl2}{rgdpl2te} \times POP.$$  

Labor $L$, which is defined as total hours worked, is the product of hours worked per worker $h$ and employment. The adult population is computed using the data from ILO of the adult share in total population and the population data from PWT.

In order to compute the household expenditure on durables $X_d$, we use the consumer expenditure data of EM and the data of PWT for consumption share of GDP ($kc$), GDP per capita ($rgdpch$) and population ($POP$):

$$X_d = \frac{\text{consumer expenditure on durables}}{\text{consumer expenditure}} \times kc \times rgdpl2 \times POP.$$  

The household income share of capital $\theta_h$ is derived from EM data on household income:

$$\theta_h = 1 - \frac{\text{gross income from employment}}{\text{gross income}},$$

### 3.2 Imputing Service Flow from Consumer Durables

Consumption expenditure $C_x$ in the data is defined as

$$C_x = C_{nd} + C_s + X_d,$$

where $C_{nd}$, $C_s$ and $X_d$ stand for the household expenditures on non-durables, services and durables. However, consumption in the model $C$ is defined as

$$C = C_{nd} + C_s + C_d,$$

where $C_d$ stands for the services flow generated from durable stocks. Investment $X$ is defined as the sum of gross domestic capital formation $X_f$ and $X_d$. Output $Y$ is defined as the sum of GDP and $C_d$. Total capital stock $K$ is the sum of net fixed capital stock $K_f$ and the stock of consumer durables $K_d$.

The service flow from consumer durables $C_d$ is imputed as

$$C_d = K_d(R_k + \delta_d).$$

where $R_k$ is the net return on capital stock and $\delta_d$ is the depreciation rate of consumer durables assumed to be equal to 0.2. The stock of consumer durables follows a law of motion:

$$K_{d,t+1} = (1 - \delta_d)K_{d,t} + X_{d,t},$$
where the stock of consumer durables in 1990 is assumed to be equal to

\[ K_{d,1990} = \frac{X_{d,1990}}{\delta_d}. \]

The net return on capital \( R_k \) is defined as

\[ R_k = \theta_f \frac{GDP}{K_f} - \delta_f, \]

where \( \theta_f \) is the income share of net fixed capital stock and \( \delta_f \) is the depreciation rate of net fixed capital stock. The income share of net fixed capital stock is derived as

\[ \theta_f = \frac{\theta_h \times NNP + \Delta}{GDP}, \]

where \( \theta_h \) is the household income share of capital which is directly obtained from data, \( \Delta \) stands for the depreciation of net fixed capital stock and \( NNP = GDP - \Delta \). The depreciation rate of net fixed capital stock is computed as

\[ \delta_f = \frac{\Delta}{K_f}. \]

Finally, total capital share \( \theta \) is defined as

\[ \theta = \frac{\theta_f \times GDP + C_d}{Y}. \]
4 Institutional and Governance Indicators - Definitions and measurement details

World Bank collects data on a set of institutional and governance indicators from 212 nations and we have the time series since 1996. In each instance, measures range from −2.5 to +2.5 with standard errors reflecting variability around the point estimate. The indicators are based on 30 aggregate data sources, survey and expert assessments. The details can be found in:


(1) Voice and Accountability - reflects perceptions of the extent to which a country’s citizens are able to participate in selecting their government, as well as freedom of expression, freedom of association, and a free media

(2) Political Stability and Absence of Violence/Terrorism - reflects perceptions of the likelihood that the government will be destabilized or overthrown by unconstitutional or violent means, including politically-motivated violence and terrorism

(3) Government Effectiveness - reflects perceptions of the quality of public services, the quality of the civil service and the degree of its independence from political pressures, the quality of policy formulation and implementation, and the credibility of the government’s commitment to such policies

(4) Regulatory Quality - reflects perceptions of the ability of the government to formulate and implement sound policies and regulations that permit and promote private sector development

(5) Rule of Law - reflects perceptions of the extent to which agents have confidence in and abide by the rules of society, and in particular the quality of contract enforcement, property rights, the police, and the courts, as well as the likelihood of crime and violence

(6) Control of Corruption - reflects perceptions of the extent to which public power is exercised for private gain, including both petty and grand forms of corruption, as well as "capture" of the state by elites and private interests.
Colonial Investments and Long-Term Development in Africa: Evidence from Ghanaian Railroads

Remi JEDWAB\textsuperscript{a} Alexander MORADI\textsuperscript{b}

\textsuperscript{a} Department of Economics, George Washington University, and STICERD, London School of Economics
\textsuperscript{b} Department of Economics, University of Sussex

This Version: November 21th, 2012

Abstract: What is the impact of colonial infrastructure investments on long-term development? We investigate this issue by looking at the effects of railroad construction on economic development in Ghana. Two railroad lines were built by the British to link the coast to mining areas and the hinterland city of Kumasi. Using panel data at a fine spatial level over one century (11x11 km grid cells in 1891-2000), we find strong effects of rail connectivity on the production of cocoa, the country’s main export commodity, and development, which we proxy by population and urban growth. First, we exploit various strategies to ensure our effects are causal: we show that pre-railroad transport costs were prohibitively high, we provide evidence that line placement was exogenous, we find no effect for a set of placebo lines, and results are robust to instrumentation and matching. Second, transportation infrastructure investments had large welfare effects for Ghanaians during the colonial period. Colonization meant both extraction and development in this context. Third, colonial railroads had a persistent impact: railroad cells are more developed today despite a complete displacement of rail by other transportation means. Physical capital accumulation and the demographic transition account for path dependence. Colonial railroads thus shaped the economic geography of Ghana.

Keywords: Colonialism; Transportation Infrastructure; Multiple Spatial Equilibria; Africa

JEL classification: F54; O55; O18; R4; F1

\textsuperscript{*}Remi Jedwab, George Washington University and STICERD, London School of Economics (e-mail: jedwab@gwu.edu). Alexander Moradi, University of Sussex (email: a.Moradi@sussex.ac.uk). We thank Robert Allen, Gareth Austin, Jean-Marie Baland, Sonia Bhalotra, Francois Bourguignon, Miriam Bruhn, Denis Cogneau, Paul Collier, Xinshen Diao, Taryn Dinkelman, Paul Dorosh, Robert Eastwood, Marcel Favohmps, James Fenske, James Foster, Richard Hornbeck, Elise Huillery, Eliana La Ferrara, Molly Lipscomb, Stephan Litschig, Rocco Macchiavello, Karen Macours, William Masters, Stelios Michalopoulos, Branko Milanovic, Nathan Nunn, Henry Overman, Thomas Piketty, Jean-Philippe Platteau, Ilia Rainer, Ariell Reshef, Elizabeth Sadoulet, Sheetal Sekhri, Stephen Smith, Francis Teal, Pedro Vicente and seminar audiences at Oxford (CSAE and Economic History seminars, CSAE Conference), Bocconi (CEPR/AMID Development Conference), London School of Economics (STICERD and Economic History seminars, SERC Conference), George Washington University, Dartmouth College (NEUDC), University of Virginia, RAND, International Food Policy Research Institute, Paris School of Economics, Carlos Tercero, Namur, Sussex, Cambridge (RES Conference), Geneva (AEHW), Kent, University of Barcelona and Stellenbosch (WEHC) for helpful comments. We are grateful to Jonathan Rothbaum for excellent research assistance. The financial support of PSE, AMID-LSE, GWU-IIEP and ESRC First Grant (RES-061-25-0456) is gratefully acknowledged.
What is the impact of colonial infrastructure investments on long-term development? An extensive literature has studied the role of colonial institutions, relating economic outcomes today to the duration of colonization (Feyrer & Sacerdote 2009), the form of colonization (extraction versus settlement: Engerman & Sokoloff 2000, Acemoglu, Johnson & Robinson 2001, 2002, Dell 2010, Bruhn & Gallego 2012; direct versus indirect rule: Banerjee & Iyer 2005, Iyer 2010) or the identity of the colonizer (Porta et al. 1998; Bertocchi & Canova 2002). Fewer studies investigated channels other than institutions. Glaeser et al. (2004) argue that human capital mattered more. Banerjee & Iyer (2005) and Dell (2010) show that colonial institutions influenced public investments but they either look at current investments or they use proxies for colonial investments such as literacy and schooling as in Dell (2010). The effect of colonial investments in physical capital has been largely ignored.¹

We look at rail construction in Africa. First, railroads were often the single largest expenditure item in colonial budgets. In Ghana, for example, railroad expenditures accounted for 31.4% of total public expenditure in 1898-1931.² Railroad construction represented two thirds of government development expenditures. Recurrent education expenditures in contrast amounted to a meagre 4%. Second, railroads are a canonical "colonial" transportation technology. 88.3% of total rail mileage in sub-Saharan Africa was built before independence. These lines served various purposes.³ Military domination, against natives or other colonial powers, was given as motivation in 35.5% of the cases, mining and cash crop agriculture were mentioned in 36.0% and 42.4% of cases respectively. The lines were often intended to support European mining companies and farmers. Yet these lines also affected the economic lives of Africans by creating economic opportunities. Because infrastructure investments are highly localized, they shaped the economic geography of colonies, with long-term effects on development.

In this paper we study the impact of rail construction on development in colonial and post-colonial Ghana. Two lines were built by the British in 1898-1918 to link the coast to mining areas and the hinterland city of Kumasi. Using panel data at a fine spatial level over one century (0.1x0.1 degrees, or 11x11 km, grid cells in 1891-2000), we find a strong causal effect of rail connectivity on the production of cocoa, the country's main export commodity, and development, which we proxy by population and urban growth.

First, we exploit various identification strategies. We document that the placement of the lines was exogenous to cocoa production and population growth. The Western Line connected the European gold mines to the coast and was extended to Kumasi for military domination. The Eastern Line connected the capital city of Accra to the network through Kumasi. First, we show that pre-railroad trade costs were prohibitively high limiting cocoa production to a narrow coastal strip. Second, we verify the parallel trend assumption using data before 1901. Third, having spatially refined data allows us to compare identical observations. Cells less than 50 km from the lines are all similar in terms of observables, yet the effects are much stronger just along the lines. Fourth, there are no effects for a set of placebo lines that were planned but not built. Fifth, as cocoa trees take five years to produce, we verify there are no effects for lines that were not built in time to affect production. Sixth, cocoa was mentioned as one of the motivations of the Eastern Line, but we claim it was an ad hoc justification. This is confirmed by the fact we find the same effects for both lines. In the event that the placement was not exogenous, we verify that our results are robust by employing two techniques used in the literature. We instrument for rail connectivity with straight lines between the two ports and the city of Kumasi, thus using the fact that being on a straight line between two large cities makes it more likely to be connected. In the spirit of nearest neighbor matching, we also only use placebo grids, those that would have been connected if the placebo lines had been built, as a control group.

Second, these investments had large welfare effects for Ghanaians during the colonial period. We find a strong effect of rail connectivity on cocoa production, population, and urban growth in 1901-1931.⁴ Donaldson (2010), Burgess & Donaldson (2010), Burgess et al. (2011) and Chaves, Engerman & Robinson (2012) also analyze colonial investments in transportation infrastructure but they do not investigate their effects on long-term development.⁵ By comparison, in Kenya in 1896-1930 the share of railroad expenditure in total public expenditure was 19.3%. In French West Africa in 1910-1956, this share amounted to 30.0%.⁶ For each rail line in Sub-Saharan Africa, we know when it was built and the main motivation behind the construction of the line. More details can be found in Web Appendix 1.
Trade costs decreased along the lines, which made cocoa production for export markets profitable. Rural population increased because cocoa cultivation required more labor in cocoa-producing villages. Urban population increased because producing villages used towns as trading stations, whether large railroad stations or small intermediary towns. We examine the welfare effects of rail construction. The social savings approach indicates that railroads account for 8.7% of GDP using cocoa only. As an alternative, we use our point estimate to show that railroads caused 44.5% of cocoa production in 1927. We find that two thirds of the surplus generated by cocoa production went to Ghanaians while the other third was captured by the colonizer. Colonization meant both extraction and development in this context.

Third, we show that these colonial investments had long-term effects on development. Railroads fell largely out of use in the 1970s, due to poor management, lack of maintenance, and competition from roads. Goods and passenger traffic collapsed after 1974, and railroads now transport one-third of what they did at independence. Moreover, Jedwab (2011) describes how cocoa production has disappeared from the old producing areas due to the shifting cultivation process characteristic of this crop. Although cells along the rail lines have lost their initial advantage in terms of both transportation and cocoa cultivation, we find that railroad cells are more urbanized, have larger manufacturing and service sectors, and better infrastructure today, showing evidence for path dependence. The dynamics of path dependence are studied using panel data on population growth in 1901-2000. We test several explanations and find that physical capital accumulation (structural change and contemporary infrastructure investments) and the demographic transition, and not sunk investments in 1931, account for this result.

Our focus on railroads also connects with the literature on transportation infrastructure, trade and development. This literature often mentions the conjunction of bad geography and poor infrastructure as the main obstacle to trade expansion in Africa (African Development Bank 2003, African Union 2006, Buys, Deichmann & Wheeler 2010). Despite this recent interest in transportation projects, little is known on their long-run effects, and more research is needed. The macroeconomic literature finds that better infrastructure diminishes trade costs and boosts exports (Radelet & Sachs 1998, Limão & Venables 2001). The micro-development literature shows that rural roads reduce poverty in connected villages by integrating labor and goods markets, thus providing new economic opportunities to their inhabitants (Jacoby 2000, Jacoby & Minten 2009, Mu & van de Walle 2011). Another strand studies the impact of large transportation projects, whether highways (Michaels 2008, Storeygard 2011) or railroads (Banerjee, Duflo & Qian 2012, Atack et al. 2010, Donaldson 2010, Burgess & Donaldson 2010, Atack & Margo 2011, Donaldson & Hornbeck 2011). They show there are significant gains from market integration.

Our findings also advance the literature on the historical roots of African underdevelopment. Gennaioli & Rainer (2007) and Michalopoulos & Papaioannou (2012) emphasize the role of pre-colonial centralization. Nunn (2008), Nunn & Wantchekon (2011) and Nunn & Puga (2012) investigate the economic impact of slave trade, while Acemoglu, Johnson & Robinson (2001), Huillery (2009) and Michalopoulos & Papaioannou (2011) study the effects of colonial rule. This paper shows how colonial infrastructure investments one century ago shaped the economic geography of Ghana. This result is in line with the literature on the existence of spatial equilibria and the role of path dependence in the location of economic activity (e.g., Davis & Weinstein 2002, Redding, Sturm & Wolf 2011, Holmes & Lee 2012, and Bleakley & Lin 2012). Similarly to Bleakley & Lin (2012) who study portage sites in the U.S., we find no evidence that railroad cells, having lost their natural advantages, are in decline.

The paper is organized as follows. Section 2 presents the historical background of rail construction, cocoa and population growth in Ghana and the data used. Section 3 explains the methodology, while Section 4 displays the main results. Section 5 discusses the welfare effects during the colonial period. Section 6 investigates why these effects persist over time, and Section 7 concludes.

---

4 In 2005, Africa had 0.002 km of railroad track per 1,000 sq km, and Ghana had around 4 km, while India had 21 km, the U.S. 25 km and Europe 51 km (World Bank 2010). Turning to roads, Africa had 85 km of roads per 1,000 sq km, and Ghana had 239 km, while India had 1,115 km, the U.S. 702 km and Europe 1,377 km.

5 Those studies are more convincing in terms of identification. Michaels (2008), Atack et al. (2010) and Banerjee, Duflo & Qian (2012) use the fact that being on a straight line between two large cities makes it more likely to be connected to a highway or a railroad. Donaldson (2010) does not find any effect for rail lines that were approved but never built.
2. BACKGROUND AND DATA

We discuss the historical background and the data we have collected for our analysis. Web Appendix 1 contains more details on how we construct the data.

2.1 New Data on Ghana, 1891-2000

In order to analyze the effect of railroad construction on development in Ghana, we construct a new data set of 2091 grid cells of 0.1¢0.1 degrees (11x11 km) from 1891 to 2000. We choose a high resolution grid since we have very precise GIS data on railroads, cocoa production, and population. We obtain the layout of rail lines in GIS from Digital Chart of the World. We then use various documents to recreate the history of rail construction. We know when each line was started and finished and when each station was reached and opened. We also located lines that were planned but not built. For each line, we create cell dummies equal to one if the Euclidean distance of the cell centroid to the line is 0-10, 10-20, 20-30, 30-40 or 40-50 km. Our main analysis focuses on the rail network in 1918. We also create a dummy equal to one if the cell contains a railroad station in 1918. We proceed similarly to construct a GIS database on transportation networks in 1901 (rivers and forest tracks) and motor roads in 1922.6

The data on cocoa land suitability was derived from maps of cocoa soils in Ghana. A cell is defined as suitable if it contains cocoa soils. It is highly suitable if more than 50% of its area consists of forest ochrosols, the best cocoa soils. It is very highly suitable if more than 50% of its area consists of first class or second class ochrosols, the best types of ochrosols. Production data was digitized from a contemporary map and we use GIS to calculate the amount produced (tons) for each cell in 1927. Production was almost zero around 1901 and we know where it was exactly located. We also have data on cocoa tonnages brought to each rail station in 1918.

The British established the Gold Coast colony in the south and extended their domination to what is now Ghana in 1896. Improving transport infrastructure was on the agenda, to permit military domination and boost trade historically constrained by high transport costs. Draft animals were not used due to the Tsetse fly transmitting trypanosomiasis. Ghana also lacked navigable waterways. Headloading was the main means of transport, although cocoa was also rolled in barrels along a few tracks. Owing to the thick primary forest in southern Ghana, there were only a few well-cleared tracks. Railroads were the transport technology of the time, but the British had to choose between a western, central or eastern route.

We use census gazetteers to reconstruct a GIS database of localities above 1,000 inhabitants. The number of these localities increased from 144 in 1891 to 2,975 in 2000.7 Since our analysis is at the cell level, we use GIS to construct the urban population for each cell-year observation. While we have exhaustive urban data, we only have georeferenced population data for Southern Ghana in 1901 and the whole territory in 1931, 1970, 1984 and 2000. We calculated rural population by subtracting urban population from total population. All cells have the same area, so population levels are equivalent to population densities.

Lastly, we have data on infrastructure provision at the gridcell level in 1901, 1931 and 2000. We also use census data on employment and human capital for each cell in 2000.

2.2 The Railroad Age

2.2.1 Railroads Built

Infrastructure investments are typically endogenous, driven by the economic potential that justifies them. Hence, a simple comparison of connected and non-connected cells is likely to overstate the output created by it. The railroad age in Ghana provides us with a natural experiment to identify the effect of reduced trade costs on development. This summary draws on Gould (1960), Tsey (1986) and Luntinen (1996). Figure 1 and Figure 2 show the geographic location of the mentioned lines.

The British established the Gold Coast colony in the south and extended their domination to what is now Ghana in 1896. Improving transport infrastructure was on the agenda, to permit military domination and boost trade historically constrained by high transport costs. Draft animals were not used due to the Tsetse fly transmitting trypanosomiasis. Ghana also lacked navigable waterways. Headloading was the main means of transport, although cocoa was also rolled in barrels along a few tracks. Owing to the thick primary forest in southern Ghana, there were only a few well-cleared tracks. Railroads were the transport technology of the time, but the British had to choose between a western, central or eastern route.

---

6 There are three types of roads: class I (roads suitable for motor traffic throughout the year), class II (roads suitable for motor traffic but occasionally closed) and class III (roads suitable for motor traffic in dry season only).

7 The number of 1,000 inhabitant localities was 144 in 1891, 143 in 1901, 353 in 1921, 437 in 1931, 627 in 1948, 1,110 in 1960, 1,262 in 1970, 1,895 in 1984, and 2,975 in 2000.
The first line followed the western route. Strong interest groups of British capitalists lobbied to connect the gold fields in the hinterland. Mines needed heavy machinery and large quantities of firewood (or coal). Headloading made gold production prohibitively costly. The colonial administration gave in to the pressure, turning down alternative lines, for which surveys attested a greater potential for agricultural exports (palm oil). Web Appendix 2 describes how the Governorship of William E. Maxwell (1895-97) was instrumental in the decision-making process. Maxwell previously worked in the Malay States where railroads served the tin mines and thought that the same model of “mining first” should be applied to Ghana. There were also military reasons to connect the Ashante capital Kumasi. The British fought four wars before they annexed the Ashante Kingdom in 1896. The railroad was meant to allow the quick dispatch of troops. Construction begun in 1898. The line started from Sekondi on the coast and reached the gold mines of Tarkwa and Obuasi in May 1901 and December 1902 respectively (see Fig. 2). The line was further extended to Kumasi in September 1903. Much of the line went through virgin forest and was opened to traffic in 1904. This line was built by Europeans for Europeans, and gold mining accounted for two thirds of the line’s traffic (in volume) in 1904-1912. Cocoa did not play any role in the location choice, but the line had a strong effect on its cultivation, as argued by Tsey (1986, p.303-306). Cocoa freight on this line increased from 0 tons in 1904 to 19,191 tons in 1915.

The second line followed the eastern route. Colonial governors long favored a central route (e.g., from Saltpond or Apam, see Fig. 2), but a series of unexpected events led to the governorship of John P. Rodger (1904-1910) who thought that the capital Accra had to be the terminus of this second line to Kumasi. By 1905, several additional motivations were cited for its construction (Tsey 1986, p.56-63): the export of palm oil, rubber, and cocoa, the exploitation of the Eastern Akim Goldfields around Kibi, and the development of tourism around Abetifi (see Fig. 2). Construction started in 1909 and the line reached Mangoase in late 1909. However, serious flooding in 1910 and 1911 meant that the line was not opened to traffic before 1912. It was extended to Koforidua in 1915 and Tafo in 1916, but Tafo station was only opened in July 1917. Rail construction had to stop due to wartime shortages, and Kumasi was connected in 1923. A potential concern is whether the placement was exogenous. Cocoa cultivation originally spread out in the Eastern province from Aburi Botanical Gardens, where the British distributed cocoa seedlings (see Fig. 2, and Section 2.3 below). The province’s production was already growing before Mangoase was reached in 1909: around 1,000 tons in 1901, 5,000 tons in 1905, 15,000 tons in 1910, 40,000 tons in 1915, 65,000 tons in 1920 and 100,000 tons in 1925. As cocoa trees take five years to produce, production before 1914 (ca. 30,000 tons) cannot be attributed to the line. Growing “in advance” could be a cause of railroad construction. For this reason, it will be important to show that: (i) transport was prohibitively costly before, so production would have remained limited to pre-railroad levels, (ii) both lines have similar effects, (iii) results are robust to controlling for the spatial diffusion of cocoa from Aburi, (iv) no positive effects are found for placebo lines, and (v) results are robust to instrumentation or matching.

2.2.2 Reduction in Transport Costs

Railroads permitted a massive decrease in transportation costs. While the freight rate per ton mile was 5s for headloading, 3.2s for canoe, 2.5s for lorries (1910, against 1s from 1925), 1.9s for cask rolling, and 1s for steam launch, it was only 0.4-0.6s for railroads. Yet this simple comparison of freight rates underestimates the magnitude of transport costs: (i) the cost above only concerns headloaders that walk along a forest track. There were only a few well-cleared tracks and headloaders often had to go through the dense tropical forest, which made it even more costly, (ii) cask rolling necessitated good quality roads and there were only a few of them then, (iii) Ghana lacked navigable waterways and these did not serve the areas where cocoa could be grown, and (iv) roads were of poor quality until 1924 when the government started the “Tarmet Program” which made roads suitable for motor traffic throughout the year (Gould 1960). Until the late 1920s, rail was by far the best transport technology.

Figure 3 displays the cells where cocoa cultivation was profitable on average using transportation networks in 1900. Web Appendices 3 and 4 describes how this map was created. Using a GIS map of networks in 1900.

---

8The line was officially opened to traffic in 1912, but it reached Mangoase three years before in 1909. There is some evidence that farmers went there as soon as 1909 to grow cocoa, expecting the railroad to be opened that year. Given that cocoa trees take five years to produce, we should not see any effect on total production before 1914.
1900 (see Fig. 3), we estimate for each cell the minimal transport cost of one ton to any port. We account for topography when calculating this cost as the topographic distance is higher than the Euclidean distance. From Cardinall (1931), we get an estimate of the production cost of one ton of cocoa, £20.9 on average. Given the coastal producer price was £31.3 on average in 1920-29, production is profitable in a cell if 31.3 - 20.9 - transport costs > 0, or transport costs < £10.4. As in Donaldson & Hornbeck (2011), the reduction in trade costs must have expanded production in the feasible region, where production was already profitable ex-ante but is now even more profitable, and in the infeasible region, where production was not profitable ex-ante but is now profitable. Figure 3 displays both regions. Pre-railroad trade costs were prohibitively high, limiting cocoa production to a narrow coastal strip. Railroads were thus essential to the colonization of the hinterland. According to Luntinen (1996, p.107), "The very existence of the transport network encouraged the production of surplus for the market. It was cocoa that made the Gold Coast the richest colony in Africa. The farmers seized the opportunity as soon as the railway reached them." From 1912 on, the share of cocoa transported by rail was around 80% (see Fig. 4).

Roads were first complementary to the railroads as they were feeders to them. The first lorry was imported in 1903, but there were only two lorries in 1914 (Luntinen 1996). Besides, roads were of poor quality until 1924 as discussed above. Roads later became serious competitors for the railroad and opened new areas to cocoa. Even if no railroad had been built, roads would have permitted the boom. Our goal is not to compare the respective impacts of railroads and roads. We focus on the railroad age because it provides us with a natural experiment to identify the effect of colonial investments on development.

2.2.3 Placebo Lines

The British had to choose between a western, central or eastern route, and various private and government lines were proposed and surveyed before the Western and Eastern Lines were built. We can address concerns regarding endogeneity by using these alternative railroad routes as a placebo check of our identification strategy. Web Appendix 2 gives a detailed background of each placebo line.

Five alternative routes were proposed before the first line was built. The aim was to ensure military domination and increase trade. Judged by observables, the proposed lines were influenced by soil quality and population density in a similar way as the actual lines built. These proposed lines all had the same probability of being built as the two lines that were indeed built, and only random events (e.g., a change in the colonial Governor) explain why construction did not go ahead. First, the Cape Coast-Kumasi line (1873) was proposed to link the capital Cape Coast to Kumasi to send troops to fight the Ashante. The project was dropped because the war ended too quickly. Second, Governor Griffith advocated the construction of a central line from Saltpond to Oda and Kumasi (1893) to tap palm oil areas and link the coast to Kumasi. When he retired in 1895, he was replaced by Governor Maxwell who favored the mining lobbies and the Western Line. Third, Maxwell also thought that the colony needed a central line. There were two competing projects with two different terminus, Apam-Oda-Kumasi (1897) and Accra-Oda-Kumasi (1897). A conference was to be held in London in 1897 to discuss the various proposals but unexpectedly Maxwell died before reaching London. Fourth, Maxwell was replaced by Governor Hodgson who favored Accra. However, he thought that the Accra line should be built to Kpong on the Volta river (1898), so as to boost palm oil and cotton production there. Construction was approved in 1903 but Governor Nathan retired in 1904 before works even begun. Fifth, Governor Rodger, who replaced Nathan, did not see any interest in a line to Kpong and he proposed the Eastern Line. Construction started in 1909.

The rail network was subsequently expanded. Hence, we also consider lines that were not built in time to affect production in 1927. Note that cocoa is a perennial crop. Pod production of the type of cocoa predominantly grown in Ghana starts after 5 years (Jedwab 2011). Hence, to observe any impact on cocoa production in 1927, farmers must have planted cocoa trees before 1922. The extension of the Eastern Line from Tafo to Kumasi (1923) is a good counterfactual for the Accra-Tafo line (1918). Another line was built

\[9\] This includes the cost of establishing a farm in the forest (five years with no production) and the annual cost of labor when trees bear fruit, assuming a farm lifetime of 30 years.

\[10\] £10.4 is an upper bound as there are various costs that we cannot account for when a subsistence farmer adopts cocoa production: risk premium due to risk aversion, losses during transport due to theft or rotting, etc.
from Huni Valley to Kade in 1927, to connect the diamond mines at Kade and encourage cocoa, kola, palm oil, and timber exports. We verify that there are no effects for these lines in 1927.

2.3 The Cash Crop Revolution and Development in Ghana

Cocoa has been the main motor of Ghana’s development, and this made it a leader of the African “cash crop revolution” (Austin 2008; Gollin, Jedwab & Vollrath 2012). Cocoa was introduced by missionaries in 1859, but it took 30 years before cocoa was widely grown, making Ghana the world’s largest exporter as early as 1911. Figure 4 shows the aggregate production and export share of cocoa from 1900-1927. Figure 5 shows grid cells that are suitable or highly suitable for cultivation and production in 1927. Cocoa originally spread out in the Eastern province from Aburi Botanical Gardens, where the British distributed cocoa seedlings (Hill 1963, p.173-176).11 As Ghanaians realized how profitable cocoa was, more and more people specialized in the crop. Why did production boom in Ashanti, around Kumasi, and not in the South-West, around Sekondi? Trade costs were lower there as it was closer to the coast. But the South-West is characterized by very poor cocoa soils and too much rainfall.12

Ghana has experienced sustained population growth after 1901. Its population increased from 1.9 million in 1901 to 3.2 million in 1931. Population growth accelerated in 1921-1931, when the annual growth rate was 3.6%, against 1.3% in 1901-1921. Population was 18.9 million in 2000. Figure 6 displays total population for southern cells in 1901, 1931, and 2000. The comparison with Figure 5 suggests that population increased in cocoa-producing areas. While Ghana was almost unurbanized at the turn of the 20th century, it is now one of the most urbanized countries in Africa. It started its urban transition earlier than most African countries, due to the cocoa boom (Jedwab 2011; Gollin, Jedwab & Vollrath 2012). Defining as urban any locality with more than 1,000 inhabitants, Ghana’s urbanization rate increased from 23.5% in 1901 to 48.6% in 1931 and 68.5% in 2000. The two largest cities are Accra, the national capital, and Kumasi, the hinterland capital. Before the 20th century, towns were state capitals or trading centres (see Dickson 1968, p.70-71). Most of the latter were on the coast, where European merchants would meet local merchants from the interior. There were also trading centers in the north, which benefitted from their location on historical trade routes. In the early 20th century, most of urban growth took place in the forest zone, with the development of cocoa production, modern transportation and mining (see Dickson 1968, p.246-261). Towns grew because they were cocoa buying centers, the homes of wealthy cocoa farmers, or market towns where they would spend their income.

3. EMPIRICAL STRATEGY IN 1901-1931

We first test if connected cells experience a boom in cocoa production, population growth, and urban growth in 1901-1931. We explain the various strategies we implement to obtain causal effects.

3.1 Main Econometric Specification

The main hypothesis we test is whether rail connectivity drives cocoa production and population growth. We follow a simple difference-in-difference strategy where we compare connected and non-connected cells over time. We run the following model for cells $c$ and years $t = [1901, 1931]$:

$$Cocoa_{c,t} = \alpha + Rail_{c,t} \beta + \gamma_t + \delta_c + u_{c,t} \tag{1}$$

where our dependent variable is the cocoa production (tons) of cell $c$ in year $t$. $Rail_{c,t}$ are cell dummies capturing rail connectivity: being 0-10, 10-20, 20-30, 30-40 or 40-50 km away from a rail line. The dummies are equal to zero in 1901. We include both cell and year fixed effects. We run a second model:

$$Pop_{c,t} = \alpha' + Rail_{c,t} \beta' + \gamma'_t + \delta'_c + v_{c,t} \tag{2}$$

11The Botanical Gardens were established in Aburi in 1890, because of its health climate and its proximity to Accra.
12The South-West consists of oxisols or intergrades, which are very poor cocoa soils. The lack of soil minerals causes low yields and premature tree aging. Annual rainfall often exceeds 2,000mm, with a very wet dry season, which favors cocoa diseases.
where our dependent variable is population of cell $c$ in year $t$. We expect rail connectivity to have a positive and significant effect on cocoa production ($\beta > 0$) and population ($\beta' > 0$). We then include $Cocoa_{c,t}$ in model (2) to see if cocoa captures the effect of railroads on population. If that is the case, it means that rail connectivity has an effect on population growth through more cocoa production along the lines. There could be an independent railroad effect on population, so our goal is not to instrument production with railroads. Our goal is just to highlight one of the mechanisms at play.

We have a panel of 2,091 cells. Our main analysis is performed on the restricted sample of suitable cells. If we use the full sample, we run the risk of comparing the southern and northern parts of Ghana. We also restrict our sample to those cells for which we have total population in 1901. We end up with 554 observations, and we believe these restrictions give more conservative estimates. Results are robust to removing the restrictions. We argue in Section 2.2 that the placement of railroads was not endogenous to production and population. We now describe the tests we perform to ensure these effects are causal.

### 3.2 Exogeneity Assumptions and Controls

First, even if the Eastern Line was endogenous, the Western Line was built for mining and military domination. Endogeneity is not a concern if we find similar effects for both lines. Second, even if the Eastern Line had been built for cocoa, production would have remained small before the line was built because trade costs were prohibitively high before. This is similar to arguing that the timing of line construction was exogenous. Third, we include controls at the cell level interacted with year dummies ($X_{c,t}$) to account for potentially contaminating factors. We control for economic activity in 1901, such as cocoa production in 1901 and through a dummy equal to one if the cell has a mine. We control for demography in 1901, by including urban and rural populations. We add physical geography variables such as the shares of suitable, highly, and very highly suitable cocoa soils, the mean and standard deviation of altitude (m), and average annual rainfall (mm) from 1900-1960. We control for economic geography by including dummies for bordering another country or the sea and Euclidean distances (km) to Accra, Kumasi, Aburi, a port in 1901, and the coast. Fourth, since we have data for 11x11 km grid cells, we compare neighboring locations that are unlikely to differ in terms of unobservables. Cell area is 122 sq km, only 40% more than Manhattan’s area. By comparison, the mean areas of Indian districts and American regions, used as spatial units in the literature, are 4,300 and 78,977 sq km respectively. Cells less than 50 km from the lines are all similar in terms of observables. If the placement is truly exogenous, the effect should strongly decrease as we move away from the line, which is exactly what we find. We can also include district-year fixed effects to control for time-variant heterogeneity. We have only 9 cells per district.

### 3.3 Parallel Trends

First, if the placement or timing is exogenous, we should observe no effect before the lines are built. We run the same model as model (2) except we consider cocoa production and urban population in 1891 and 1901. We have no data on rural population in 1891. There are no significant effects in 1891-1901, while we will show there are strong effects in 1901-1931.

### 3.4 Placebo Regressions

As explained in Section 2.2.3, five lines were planned but never built. Two lines were built after 1923. For each placebo line, we create a placebo treatment dummy equal to one if the cell is less than X km from the line. First, we expect no effect for the placebo cells. Second, one issue here is that some of the placebo lines intersect with the area of influence (e.g., 0-20 km) of the two existing lines, so that there might be a correlation between the treatment and placebo dummies. Therefore, we verify that there are no effects for

---

13As explained in Section 2.1, we only have georeferenced population data for Southern Ghana in 1901.

14There were five mines in 1931: three gold mines, one diamond field, and one manganese mine. Mining exports amounted to 24.2% of exports and the number of Africans engaged in mining was 12,048. Cocoa and mining accounted for 94.5% of exports.

15We regress each control on the rail dummies using the 40-50 km cells as the omitted group. There are no significant differences, except for rural population and having a mine for the 0-10 km cells, but these effects are small.

16The coefficients (p-values) of the 0-10, 10-20, 20-30, 30-40 and 40-50 km dummies in 1891-1901 are: -2 (0.62), 14 (0.37), -6 (0.51), 3 (0.50) and 5 (0.40) for tons of cocoa production, and 90 (0.76), 40 (0.88) and -155 (0.63), 70 (0.71) and 9 (0.96) for urban growth.
the segments of these lines that do not intersect with existing lines. Third, we also compare the placebo cells with the other control cells, by dropping the railroad cells.

3.5 Instrumentation

We instrument the treatment with the distance from the straight lines between the two main ports, Sekondi and Accra, and the hinterland city of Kumasi. This strategy echoes the works of Michaels (2008) and Banerjee, Duflo & Qian (2012), who respectively instrument U.S. highways and Chinese railroads with the distance from the straight line joining two major cities, exploiting the fact that transportation networks tend to connect large cities. The instrument is valid as long as it is not correlated with any uncontrolled variable that affects the outcome. The Western Line linked Sekondi to the mines of Tarkwa and Obuasi and was extended to Kumasi for military domination. It went through dense tropical forest and the random location of the mines explained why this interior line was built from Sekondi to Kumasi. Regarding the Eastern Line, Accra was the administrative and economic capital of Southern Ghana while Kumasi performed the same role for the hinterland. It was obvious that the two largest cities would be connected at one point. We also use as an instrument the distance from the straight lines Sekondi-Tarkwa-Obuasi and Accra-Kibi (see Fig. 3) to exploit the fact that mining was a major motivation behind rail building.

3.6 Matching

We compare the connected cells with cells that would have been connected if the placebo lines had been built. This guarantees that treatment and control cells are similar in terms of unobservables. We regress each control on a dummy equal to one if the cell is less than 20 km from a 1918 rail line. We alternatively consider as a control group all the suitable cells (see Col. (1) of Table 1), placebo cells (see Col. (2)) and placebo cells using Cape Coast-Kumasi only (see Col. (3)). Placebo cells are similarly defined as cells that are less than 20 km from a placebo line. First, when compared to all suitable cells, treated cells have a larger rural population and are closer to main cities, which could lead to an upward bias, and have worse cocoa soil quality, which could lead to a downward bias. It is not obvious in which direction coefficients would be biased. Second, the same biases exist when comparing treated cells and all placebo cells. Third, we can compare treated cells with each placebo line, as some of them could prove a better counterfactual. For example, when compared to cells along the placebo Cape Coast-Kumasi line (1873), treated cell are worse (soil quality, altitude and distance to Accra or the coast) or similar across all dimensions. Using these cells as a control group should lead to a downward bias and give more conservative estimates. Results are the same whichever control group is selected.

4. RESULTS

In this section, we display the main results and examine their robustness.

4.1 Main Results

Table 2 shows the main results for cocoa production and population growth. Column 1 reports the results for model (1), while columns (2)-(13) display the results for model (2). All regressions include cell and year fixed effects and controls. We find a strong effect of rail connectivity on production, but this effect decreases as we move away from the rail line and is zero after 40 km (see Col. (1)). There is a strong effect on population growth up to 20 km (see Col. (2)). People tend to live in the vicinity of the line, although there is some production beyond 20 km. Interestingly, the rail effect is lower when we include cocoa production, which has a strong effect on population growth (see Col. (3)). The remaining rail effect is picked up by the cell dummy for having a railroad station in 1918 (see Col. (4)), but that effect also becomes lower and non-significant when we include the amount brought to the station in 1918 (see Col. (5)). This means the rail lines have a strong effect on population growth, and that this growth is coming from opportunities in the cocoa sector and other sectors if there are intersectoral linkages. Treated cells have less rainfall but the difference is small: 4% less than the average of 1477 mm. Including population in model (1) does not change the railroad effects on production. This confirms that the relationship is not from railroads to population and then to production. Jedwab (2011) discusses how the causality is unlikely to run from...
The railroads have two effects on population growth, which we study by looking at rural growth and urban growth. We call the first effect, the number of additional inhabitants per ton of cocoa produced, the labor effect \(1.53^{***}\), see Col. (5)), as more cocoa production requires more labor. The comparison of Columns (5), (9) and (13) indicates that most of the labor effect takes place in villages \(1.15^{***}\), see Col. (9)). This is logical as cocoa is produced on farms surrounding producing villages (Jedwab 2011). We call the second effect, the number of inhabitants per ton of cocoa transported, the trade effect \(0.86^*\), see Col. (5)), as more cocoa being transported requires larger rail stations. The comparison of Columns (5), (9) and (13) shows that the trade effect only occurs in towns \(0.91^{**}\), see Col. (13)). When using the urbanization rate as the main outcome, we find positive effects but only significant at 15% (not shown, but available upon request). Indeed, rural population increases almost as much as urban population.

Do the effects vary for various size categories of towns: 1,000-2,000 (small), 2,000-5,000 (medium) and 5,000-10,000 (large)? Running the same model as in Col. (10), we find no effect for small towns, a small effect for medium towns and a strong effect for large towns. If we run the same model as in Col. (13), the labor effect is mostly explained by medium towns, and many of them did not exist in 1901. The trade effect is mostly explained by large towns, and many of them already existed in 1901. To sum up, the railroads induced a cocoa boom, which had various effects on population growth: more producing villages up to 30 km away, more small producing towns up to 20 km and larger rail stations up to 10 km away.

4.2 Alternative Identification Strategies

Table 3 displays the results when we implement the various identification strategies. Column (1) replicates our main results from Table 2. For the sake of simplicity, we only focus on the 0-20 and 20-40 km dummies for production and the 0-20 km dummy for population, as there are no effects beyond.

**Western Line vs. Eastern Line.** Column (2) of Table 3 shows there are stronger effects for highly suitable cells. When comparing the results for the Western and Eastern Lines (see Col. (3) and (4)), we find lower effects for the Western Line, but this is explained by the fact that it goes through poorly suitable cells. If we restrict the analysis to highly suitable cells only, the effects are not significantly different for both lines.

**District-Year Fixed Effects.** Column (5) of Table 3 shows results are robust when we include district-year fixed effects. Using the district boundaries in 2000, we have 554 cells in 62 districts, or 9 cells per district.

**Placebo Regressions.** Table 4 displays the results of the placebo regressions. Panel A, B and C show the results for production, while Panel D, E and F show the results for population. For the sake of simplicity, we only use 0-20 km dummies, so we test whether there are positive effects just along the placebo lines. The only concern is cocoa production for Accra-Oda-Kumasi 1897 (see Col.(4)). However, this effect exists because the line intersects with the area of influence of the Eastern Line. We find no effects for the segments of the placebo lines that do not intersect with the existing rail lines, as shown in Panel B and E. The issue here is that the railroad cells belong to the control group, which leads to significant negative placebo effects. In Panels C and F, we drop the railroad cells to conduct our preferred placebo test, which only compares the placebo cells and the other control cells. Except for Accra-Kpong 1898, coefficients are very small and never significant.

**Matching.** Columns (6) and (7) of Table 3 show that coefficients are even higher when all the placebo cells or just the cells along the Cape Coast-Kumasi placebo line (< 20 km) are used as the control cells. Results are robust to using the other placebo lines (not shown, but available upon request). Results are also robust if we restrict the control group to the cells in the 40-50 km range.

**Instrumentation.** In Column (8), we instrument the rail dummies by dummies for being 0-20 and/or 20-40 km from the straight lines Sekondi-Kumasi and Accra-Kumasi. In both cases, the IV F-statistic is strong enough (5.7 in Panel A and 63.7 in Panel B), and results are unchanged. Results are the same if we use population to cocoa in 1931. First, settlement was limited in tropical forests due to thick vegetation, high humidity and disease incidence. Farmers overcame these constraints when they could grow cocoa. Second, cocoa cultivation did not depend on cities for the provision of capital and inputs, as it only required forested land, axes, machetes, hoes, cocoa beans and labor. This traditional mode of production was not conducive to a role for cities in the diffusion of innovations.
the distance from the straight lines Sekondi-Tarkwa-Obuasi and Accra-Kibi as the instrument instead.\textsuperscript{19}

### 4.3 Robustness and General Equilibrium Effects

We now perform a few robustness checks and discuss general equilibrium effects.

**Robustness.** Results are similar if we use the Euclidean distance to rail stations instead (see Col.(9)), if we drop the controls (see Col. 10)), if we drop the nodes of the network (see Col. (11)), if we use the full sample (see Col. 12) and if we use a log-linear functional form (see Col. (12)). We have no measure of total population for the full sample in 1901, but the results are the same for the urban population (not shown, but available upon request). Second, results are robust if we control for trade costs given transportation networks in 1901 and/or the road network in 1922 (not shown, but available upon request). We use roads in 1922 because they could affect production 5 years later. Estimates are slightly higher if we restrict our sample to the infeasible region. By definition, production is already profitable on average in the feasible region. We expect lower effects, unless there is land and farmer heterogeneity and the decrease in trade costs has an extensive margin effect in this region as well. Accordingly, we find a strong effect in the 0-10 km range (+595.5\textsuperscript{**}), and no significant effect beyond. Third, we may underestimate standard errors if there is spatial autocorrelation. Results are robust to having Conley standard errors using a distance threshold of 50 km or more, or clustering them at an aggregate spatial level, such as provinces in 1931 or 2000 (not shown, but available upon request).

**Magnitude and the Reversal of Fortune.** We estimate that the railroad effects explain 44.5\% of the change in cocoa production in 1901-1927 and 46.5\% of population growth in 1901-1931. Another way to assess their magnitude is to test if they were large enough to produce a reversal of fortune in Southern Ghana. We run the same model as before except we add a second treatment variable, a dummy equal to one if the year is 1931 and the cell already had a town in 1891. We thus compare the railroad effects with the effect of being historically developed, using other cells as a control group. The effect of pre-railroad development is zero for production and slightly lower than the railroad effects for total population (not shown, but available upon request). In 1901, railroad cells were half as populated as the already developed cells. In 1931, there is no significant difference anymore. No reversal of fortune is observed, but railroads allowed some cells to entirely catch up with the most developed cells.

**Results on Infrastructure.** For each cell, we know the number of government and non-government schools and European and African hospitals, and whether the cell was crossed by a class 1, class 2 or class 3 road in 1901 and 1931. We use the same regression as model (2) to test whether railroad cells got better infrastructure over time, although no significant difference is observed in 1901. There are strong positive effects on the number non-government schools (+0.70\textsuperscript{*} for 0-10 km, given a mean of 0.22) and African hospitals (+0.13\textsuperscript{*} for 0-10 km, given a mean of 0.01), and the probability of being crossed by a class 1 or class 2 road (+0.20\textsuperscript{*}, +0.29\textsuperscript{***} and +0.22\textsuperscript{**} for 0-10, 10-20 and 20-30 km respectively, given a mean of 0.24). We find no significant effects for European public goods, whether government schools or European hospitals. These effects strongly decrease when we control for population. This confirms that railroads increased population density, and public goods had to be created as a result.

**Results on Height.** We have data on all Ghanaian male recruits in the British Army in 1888-1960. We know their year and place of birth, the year they enlisted, and their height. We restrict our analysis to soldiers that served during World War I or later. In total, we have height data for 5,447 soldiers across 298 cells. Using a regression similar to model (2), with the same controls plus district-year fixed effects, and individual controls for age, farming, skills, and ethnicity, we compare soldiers born in railroad vs. non-railroad cells (using 40 km as a cut-off) before and after 1923. Better standards of living should increase height for native soldiers born just after the cocoa boom. Railroad cells start with a disadvantage (-0.63 cm), which they more than overcome over time (+1.70\textsuperscript{*} cm). This height improvement is four times the

\textsuperscript{19}Another advantage of instrumentation is that it solves the classical measurement error problem. In the presence of non-classical measurement errors, the IV estimator is biased upward. Measurement errors are endogenous if production was better measured along the lines and was underestimated beyond 40 km. It is not a concern here as total production was 218,200 tons in the 1927 map, against 210,600 tons that were registered at the ports for that year. We use exhaustive census data for population.
improvement for the sample as a whole over the same period (+0.38 cm).\textsuperscript{20} This effect is halved and becomes not significant when we control for cocoa production in 1927, which becomes significant. This confirms that the railroad effect goes through more production and better standards of living.

**Migration vs. Natural Increase.** The total population of the 554 cells doubled from 1901-1931. This growth was mostly due to in-migration from the non-forested areas. Hill (1963) describes how the "migration involved individual Akwapim, Krobo, Shai, Ga and other Ghanaiian farmers from south of the forest belt, in buying forest land which, at the time of purchase, was hardly inhabited." Here is a simple model of population growth for cell $c$ at time $t$: $\Delta P_{c,t} = (CRB_{c,t-1} - CRD_{c,t-1})P_{c,t-1} + (I_{c,t-1} - E_{c,t-1})$. Population growth $\Delta P_{c,t}$ can either be explained by natural increase (when the crude rate of birth $CRB$ is higher than the crude rate of death $CRD$), or by migration (when there are more immigrants $I$ than emigrants $E$). First, if growth was driven by natural increase, we would expect a strong effect of $P_{c,t-1}$ on $\Delta P_{c,t}$. We control for 1901 population in all regressions. Since the results are the same when initial population is not included as a control, natural increase did not play a major role here. Second, until 1921, the birth and death rates were similar, around 52 per 1,000 people, and natural increase was small. In 1921, the death rate decreased to 25 in the main towns and 44 elsewhere. Natural increase contributed to population growth from 1921, but this effect was larger for towns. Thus, migration was the main source of population growth, and this migration effect was reinforced by natural increase in the 1920s.

**Migration and General Equilibrium Effects.** If railroads reallocate labor across space, does overall welfare increase? First, are control cells negatively affected by migration? We verify that non-railroad cells that were historically populated gained population between 1901 and 1931 when compared to other control cells. Locations that already had a city in 1891 grew further after 1901, and most placebo cells did not lose population when compared to other control cells (see Panel F of Table 4). Yet it is likely that some cells of the non-forested areas, which are not included in the sample of suitable cells, lost population and welfare may have decreased as a result. Second, has overall welfare increased despite the loss of population in some areas? People migrate because they obtain a higher utility at the destination location (Harris & Todaro 1970). Migration is rational and it leads to a more efficient spatial allocation of resources. In Ghana, the railroads gave access to a new factor of production that made people more productive, forested land that could be used to grow cocoa for export. Using data on production and trade costs for cocoa farmers in 1930, we find that cocoa farmers are 45-90% wealthier on average than subsistence farmers, including home production.\textsuperscript{21} Subsistence farmers accounted for almost 95% of employment in 1901. But the employment share in the cocoa sector increased from almost 0% in 1901 to one third in 1931.\textsuperscript{22} In other words, one third of the population became more productive because they gained access to a new factor of production. Our framework is thus different from other studies of infrastructure projects, for which selective migration could be an issue (see Dinkelman 2011 for an interesting discussion in the case of electrification). This allowed Ghana to become one of the wealthiest African countries (Austin 2008; Gollin, Jedwab & Vollrath 2012). Ghana was the fourth most urbanized Sub-Saharan African country in 1931 and had the fifth highest per capita GDP in 1950. More than half of population growth happened in cities. Railroads thus caused a qualitative change in population, by increasing overall population density. We also find positive effects on infrastructure and height.

5. DISCUSSION ON RAILROADS AND DEVELOPMENT DURING THE COLONIAL PERIOD

In this section we offer a general discussion on the role of railroads in development during the colonial period. First, we study the channels through which railroads transformed the colonial economy. Second, we estimate the welfare implications of railroad construction.

\textsuperscript{20}If we compare soldiers born before and after 1914, thus also including the potential effects of better standards of living on these soldiers who were already 1-9 year-old when cocoa production was booming, the effect increases to +2.12 cm.

\textsuperscript{21}Web Appendix 4 describes how this income gap was estimated.

\textsuperscript{22}The value of total exports per capita was multiplied by 8 during that period, around $297 in 1927 against $36 in 1901 (2010 USD). Cocoa explains the export boom. While production was almost zero in 1901, it amounted to 80% of exports by 1927. Using aggregate and export data, we verify that the production of other cash crops was unaffected by the cocoa boom.
5.1 Colonial Railroads, Trade, and Urbanization

Here is the conceptual framework we have in mind. Assume a poor country is divided into grid cells. Each cell lives in autarky, and productivity is such that everyone lives at the subsistence level. Most cells are a tropical forest with limited settlement. Cash crops could be produced there but there are no means of transportation. For exogenous reasons, some cells are connected to the rest of the world via a new transportation network. They specialize in their comparative advantage and export cash crops against food and non-food imports. Non-connected cells remain in autarky and close to the subsistence level. In this framework, towns grow as intermediary trading stations between producing areas and the ports. We use data on trade and urban employment to provide more evidence on these channels.

Colonial Railroads and Trade Integration. First, we find that trade (exports plus imports) accounted for 74% of railroad traffic (in volume) on average in 1904-1931. More precisely, it was 35% for exports, 39% for imports and 26% for internal traffic. In the period 1904-1918, imports accounted for 58%, because the Western Line was used to import building materials, machinery and coal for the gold mines. In the later period 1919-1931, exports accounted for 55%, because cocoa boomed along both lines, as well as manganese and timber exports. Second, we can easily link the export and import structures and railroad traffic. The export structure was the following in 1930: cocoa 76%, mining 23% and timber 1%. The traffic structure for export was: cocoa 72%, mining 24% and timber 1%. The import structure was the following in 1930: food, drinks, and tobacco 30%, clothing 20%, other consumption goods 15%, construction 9%, fuels 4%, machinery 6%, and transport equipment 10%. We find a rather similar traffic structure for import: food, drinks, and tobacco 31%, clothing 11%, construction 12%, fuels 16%, machinery 9%, and transport equipment 6%. This analysis confirms the role of railroads in promoting trade.

Colonial Railroads and Urbanization Why do railroads increase urbanization? First, many farmers settle in towns as they offer better conditions of living. Using data from the 1931 census, we find that 48.5% of the urban male workforce works in agriculture. The data does not distinguish cocoa and other farmers, but this figure shows these farmers probably work on farms outside the city limits. Second, towns serve as trading stations for cocoa exports (transportation to the coast) and imports (from the coast). Trade accounts for 20.6% of urban male employment, with 15.2% being traders, 1.2% being cocoa brokers, and 4.2% working in the transport sector. Railroad workers contributed to less than 0.75% of urban change in the south between 1901 and 1931. Third, cocoa farmers are wealthier than subsistence farmers. Cocoa generated a surplus for these farmers, and the Engel curve implies they spent more on "urban" goods and services, i.e. goods that were produced in the cities or distributed through the cities (Jedwab 2011; Gollin, Jedwab & Vollrath 2012). The analysis above shows that food, drinks, and tobacco, clothing, and other consumption goods amount to two thirds of imports. Some of these goods were also produced locally, as we find that light manufacturing and services account for 30.9% of urban male employment: domestic servants (6.6%), carpenters (4.5%), tailors (2.4%), government civil servants (2.3%), masons (2.0%), goldsmiths (1.3%), policemen (1.3%), teachers (1.1%), and washermen (1.0%).

Colonial Railroads, Trade and the Central Place Theory Our results are in line with the central place theory (Christaller 1933): Settlements are central places that provide services to surrounding areas and urban systems hierarchically organize these settlements. We provide evidence for the emergence of a complex urban system as a result of colonial investments. First, villages are where cocoa beans are produced. Second, small-sized towns serve as relay stations between the producing villages and the railroad stations. Third, medium-sized towns are railroad stations from which cocoa beans are brought to the coastal ports. Fourth, large cities are ports through which cocoa beans are exported or administrative centers.

5.2 Welfare Effects

In this section we contrast the welfare gains from railroad construction based on cocoa production alone using: (i) the social savings approach, and (ii) a more direct approach estimating producer rents.

5.2.1 Social Savings Approach

The social savings are calculated as the cost difference between railroads and the next-best transportation alternative $a$ (Fogel 1964): Social savings = $(c_a - c_r) \times R$, where $c$ is the marginal cost of the transportation technology and $R$ is the total volume transported by rail. Thus, social savings are the savings to society if
the goods are transported using the new, lower-cost technology. Details of the social savings calculation are provided in Web Appendix 5. We consider head porterage as the main alternative to railroads.\textsuperscript{23} The railroad transported a total of 13.5 million ton miles of cocoa at costs of about £350,000 in 1927. If the same volume had been moved by headloading instead, the hypothetical costs would have been £3.7 million. The cost difference are the social savings, i.e. 8.7% of GDP\textsuperscript{24} We can also relate the benefits to the cost of railroad construction. Cocoa represented only 30% of freight revenues, though rail building required the same fixed capital. In 1927, the railroad’s capital outlay was £8.5 million, implying a social rate of return of 39.6% for cocoa alone. This social rate of return is impressive. For comparison, the social rate of return to primary education is below 20% (Psacharopoulos & Patrinos 2004).

Under perfect competition and inelastic demand for transport, social savings are identical to the gain in consumer surplus brought about by railroads (see Web Appendix 5). If these conditions do not hold, social savings are larger than the increase in consumer surplus (Fogel 1979). First, demand was highly elastic in Ghana, as railroad construction triggered significant cocoa production inland. Without railroads, production would have been 44.5% lower. Second, railroads were much cheaper than available alternatives in 1901. Hence, for Ghana, social savings and actual welfare gains may differ considerably.

5.2.2 Producer Rents

Our GIS analysis allows us to estimate the producer rents created by the railroad as follows: $\Delta$Producer rent $= \sum_i \left[ \pi_{1918,i} - \max(\pi_{1900,i}, 0) \right]$, where $\pi$ represents the profit of growing cocoa at observed location $i$ in 1927. $\pi_{1900}$ are profits based on the 1900 transportation costs, whereas $\pi_{1918}$ refers to 1900 transportation plus the rail network of 1918 (see Web Appendix 5). We assumed costs of a small "native" cocoa farm, and we took the 1920-29 average producer price as a basis for revenues. For 88,900 tons or 40.7% of cocoa production in 1927, $\pi_{1900}$ is negative. In these cases, cocoa production is unprofitable, cocoa would not have been grown and therefore the rent is zero. With the 1922 rail network only 3.5% of cocoa production is calculated as unprofitable - which is within the error margin of heterogeneity in production costs. Overall, we find that railroads increased producer rents by £1,264,895 or 3.3% of GDP. The real contribution to GDP is much higher as production costs and transportation also create added value.

Both approaches ignore externalities. Chaves, Engerman & Robinson (2012) pointed to the ending of slavery as one positive externality. Austin (2005) attributed the ending of slavery in Ashante to the adoption of cocoa cultivation. Our evidence suggests this connection may have existed between railroads and slavery. Without railroads, there would have been less cocoa, and a slower decline in slavery. In our paper, we tested for the existence of other externalities, increasing population densities and urbanization.

5.2.3 Rent Extraction

The colonial administration and European cocoa buying firms captured parts of the rail-induced cocoa rent. In 1920-1929, cocoa farmers received on average 79% of the export price. The price wedge results in a sum of £1,635,226, of which only 44.5% can be attributed to the railroad. In sum, while Ghanaians received rents of 3.3% of GDP, Europeans received 1.9% of GDP. Railroads generated rents on the order of 5.2% and two thirds of the surplus went to Ghanaians. Colonization meant both extraction and development in this context. The issue here is why Europeans did not capture the whole rent. First, Europeans could not produce cocoa themselves as the mortality rate of settlers was high in the tropical forests of West Africa. Second, the colonial strategy in Ghana was both extractive and non-extractive. Europeans also thought that they were bringing civilization, whether christianity or capitalism (through trade). For example, Luntinen (1996, p.57) writes: "The proposed railway was believed to be beneficial for the natives of the Gold Coast. Of course, the main objective of the several schemes was to advance trade and industry on the coast, to develop the colony, and to bring profit and revenue, i.e. to make the British colonialists rich and British imperialists mighty. But this did not exclude a benevolent attitude towards the natives."

\textsuperscript{23}With the expansion of roads from the 1920s on, the next-best alternative was lorry transport. In this exercise, however, we do not consider motor transport as counterfactual, but rather as part of modern transport technology introduced in Ghana.

\textsuperscript{24}Applying the costs of 5s per ton mile to all goods would give social savings of 27% of GDP. Fogel (1964) found social savings of 5% for the U.S. in 1890. Donaldson (2010) found social savings of 9.7% for India.
6. COLONIAL RAILROADS AND LONG-TERM DEVELOPMENT IN GHANA

In this section we document the decline of colonial railroads in modern Ghana and study their effects on long-run development. Railroad cells are more developed today, although they have lost their initial advantage in terms of transportation and cocoa cultivation. We investigate the channels of path dependence.

6.1 Quantitative Evidence on the Decline of Colonial Railroads

By 1931, 500 miles of track had been laid, and rail transported 759,000 tons of goods (1,518 tons per mile) and 1,336,000 passengers (2,673 passengers per mile). The network reached its maximum size, 750 miles, in 1957 the year Ghana became independent. From 1944 to 1974, rail transported on average 2,400 tons of goods and 8,000 passengers per mile. Traffic collapsed after 1974. In 1984, rail only transported 500 tons of goods and 2,900 passengers per mile. Traffic never recovered, and rail only transported 970 tons of goods and 2,900 passengers per mile in 2000. Similarly, while railroads accounted for more than 70% of cocoa transport until 1970, this share decreased to one third in the 1980s and 7% in 2000. Rail only transports manganese and bauxite now, these commodities being too bulky for road transport.

What caused the obsolescence of rail in the late 1970s? Luntinen (1996) describes how considerable road investments and underinvestments and management issues in the rail sector produced a significant decline of the latter. The first government of independent Ghana massively invested in roads, which were deemed more modern and three times cheaper to build. However it was argued that maintenance costs were much lower for railroads, which were more competitive in the long-run. Second, political and economic instability had a damaging effect on public investments. By 1980, track, motive power, and rolling stock were in desperate physical condition. Third, there were management issues. In 1974, the Ghana Railway Corporation (GRC) employed a staff of 15,000, twice as many as in 1958 although traffic had not increased. Payroll absorbed 70% of rail expenditure and the GRC had been in deficit since 1966.

Moreover, an agronomic feature of cocoa is that it is produced by "consuming" the forest. Cocoa farmers go to a patch of virgin forest and replace forest trees with cocoa trees. Pod production peaks after 25 years, and declines thereafter. When cocoa trees are too old, cocoa farmers start a new cycle in a new forest. Removing forest trees alters the original environmental conditions and replanted cocoa trees die or are much less productive. Jedwab (2011) uses district panel data from 1901-2010 to describe how cocoa production has disappeared from the original regions of cultivation. Production density in the Eastern province, along the Eastern Line, peaked in 1938 (12.9 tons per sq km of forested land) and continuously decreased afterwards (4.4 on average in 1960-2000). Production density in the Ashanti province, along the Western Line, peaked in 1964 (12.1) before continuously decreasing afterwards (3.9 on average in 1980-2000). Railroad cells have lost their initial advantage in both transportation and cocoa cultivation. In these conditions, should we also expect these locations to be in decline?

6.2 Evidence on Colonial Railroads and Path Dependence

We use various development outcomes to test whether railroad cells remain more developed today. We run this cross-sectional regression for suitable cells $c$ ($N = 554$) in year 2000:

$$Y_{c,2000} = \alpha'' + Rail_{c,1918}\beta'' + X_c\zeta + w_c$$

(3)

with $Y_{c,2000}$ being a development outcome in 2000, $Rail_{c,1918}$ the set of rail dummies (using the lines in 1918) and $X_c$ the same controls as before. Regression results for model (3) are reported in Table 5. In Columns (1)-(2), we show that railroad cells are more urbanized, whether we drop railroad nodes or not. When adding a dummy for having being historically developed (having a town in 1891), the railroad effect is twice as large as the pre-railroad development effect, which is not significant anymore, indicating a reversal of fortune (not shown, but available upon request). In Column (3), we use satellite data on night lights as an alternative measure of development, as in Storeygard (2011) and Henderson,

25This instability includes the overthrow of Nkrumah in 1966, the succession of military coups after 1966, the economic downturn in 1966-1969, and the economic crisis in 1974-1983
Railroad cells have a higher share of inhabitants living in residences with solid walls (see Col. (3)), higher employment shares in manufacturing and services (see Col. (5)-(7)), and better access to communication, education and health infrastructure, as well as higher levels of human capital. These effects decrease as we move further away from the line, they seem stronger for more exclusive public goods (those with a lower mean) and are reduced when we control for population density and infrastructure provision in 1931. Even if these cells are more developed today, there may have been convergence after 1931. Has the level effect narrowed over time? As in Bleakley & Lin (2012), we use population panel data over one century to study the dynamics of path persistence. We run the following model for 554 suitable cells $c$ and years $t = [1901, 1931, 1970, 2000]$:

$$
\text{LogPop}_{c,t} = \text{Rail}_c \beta_t + \gamma_t + \delta_t + X_c \zeta_t + v_{c,t} 
$$

with $\text{LogPop}_{c,t}$ being the log of population and $\text{Rail}_c$ the rail dummies. We use a log-linear model as Ghana’s population boomed after 1931, and taking logs removes any linear trend. For each year $t$, $\beta_t$ are the estimates of the rail effects relative to 1901. We include cell and year fixed effects, and the same controls as before interacted with year dummies. Figure 7 displays the effects $\beta_t$. While total and rural populations boomed in 1901-1931, the level effect is much lower (but still significant) in 2000. Urbanization has followed a different pattern, as the level effect in 1931 was halved in 1970 before increasing even further in 2000. Since we have urban data every decade, we can thoroughly study urban patterns. The level effect was high in 1931, decreased in 1948, increased in 1960, was halved in 1970 and 1984, and increased again in 2000. These patterns are verified whether we increase the urban threshold to 5,000 or use the urbanization rate as the main outcome (not shown, but available upon request). Interestingly, while the railroad effects are only significant at 15% in 1931, they become highly significant in 2000. This confirms that these locations became increasingly urban over time.

A look at the history of Ghana helps us understand these patterns. Aggregate cocoa production was halved between 1938 and 1944, because of the cocoa swollen-shoot virus disease. Hill (1963) describes how the cocoa towns were affected by the shock. In 1960-1984, the political and economic crisis, the movement of cocoa cultivation to other areas, and the decline of railroads could explain why the towns of the old producing regions were relatively "losing". Between 1966 and 1984, per capita GDP decreased by 30% and the urban sector was hit hardest (Jedwab & Osei 2012). Growth resumed after 1984, and productivity and employment rose in the urban sector. It seems railroad cells had an initial advantage but this advantage only persists if the economic situation makes it possible.

### 6.3 The Channels of Path Dependence

We study how the rail effects on urban population today vary as we control for the various potential channels of path dependence. We run the following model for 554 suitable cells $c$ and year 2000:

$$
\text{LogU Pop}_{c,2000} = \text{Rail}_{c,1918} \beta + \theta \text{LogU Pop}_{c,1901} + \zeta \text{LogRPop}_{c,1901} + X_c \lambda + \mu_c 
$$

with $\text{LogU Pop}_{c,2000}$ being the log of urban population and $\text{Rail}_{c}$ the rail dummies. $\text{LogU Pop}_{c,1901}$ and $\text{LogRPop}_{c,1901}$ control for urban and rural population in 1901. $X_c$ is the same set of controls as before. We drop the cells including the nodes. Results are reported in Table 6. Column (1) shows that railroad cells are more urbanized today. These effects disappear when we control for urban and rural population in 1931 (see Col. (2)). The long-run effects are thus explained by the level effect in 1931.

What could explain the effect of population density in 1931 on population density in 2000? Bleakley & Lin (2012) contrast the respective roles of historical and contemporary factors. First, *sunk investments* in 1931 could induce people to stay at these locations. If schools and hospitals are expensive to build, people are less mobile and original advantages have long-run effects. Second, if there are returns-to-scale in production, factors need to be co-located in the same locations. There is a *coordination problem* as it is not obvious which locations should have the contemporary factors. In this case, it makes sense to locate factors in locations that are already developed, for example the railroad cells. Third, most papers...
on multiple spatial equilibria have focused on developed countries, while this study looks at an African country that is experiencing a demographic transition. The population of Ghana was five times larger in 2000 than in 1931, for the mere reason that the death rate dropped after 1945. In this context, we should expect any original advantage to be mechanically "multiplied", although this does not tell us whether rising population densities will result in higher productivity. We use data on investments in 1931 and 2000 and population panel data from 1931-2000 to examine the respective roles of sunk investments, coordination failures, and the demographic transition in explaining path persistence.

First, the 1931 population effects are unchanged when controlling for sunk investments (see Col. (3)). Public goods in 1931 have positive effects on urban density in 2000 but they do not explain path persistence. Second, the 1931 population effects decrease when we control for the cell population history, i.e. urban and rural population in 1970 and 1984. The population effects are much stronger in 1984. In Columns (6), we examine how these effects change when we control for contemporary factors: physical capital, which we capture with manufacturing and service employment (i.e., structural change), human capital, which we capture with the literacy rate and the share of adults with secondary education, and public capital, which we capture with various measures of infrastructure provision. These variables are the same as in Table 5. Human capital does not modify the population effects (not shown, but available upon request). Controlling for structural change and current infrastructure reduce them (see Col. (5)). Yet the 1984 population effects are still large and significant. These effects cannot be explained by coordination failures. We attribute them to demographic growth. Given our results, if urban (rural) population is 100% higher in one cell in 1984, it is 142% (133%) higher in 2000. We then split the urban sample into 1,000-5,000 localities and 5,000-+ localities. We find that coordination failures matter more for 5,000-+ localities as the effects of contemporary factors (structural change and infrastructure provision) strongly increase (see Col. (6) vs. Col. (5)). Demographic growth also matters as localities above 5,000 inhabitants grow further after 1984 (see Col. (7)). The coordination problem matters less for smaller towns (see Col. (8)), whose growth is a function of rural population and urban population in 1,000-5,000 localities in 1984 (see Col. (9)). These demographic effects seem large, but they match results using aggregate data. In 1984, the crude rates of natural increase were 31 and 23 per 1,000 people in rural and urban Ghana. They imply that if urban (rural) population is 100% higher in one cell in 1984, it is 144% (163%) higher in 2000. Rural migration to the cells in which contemporary factors are co-located could explain why we find higher urban growth rates in localities above 5,000 inhabitants (153%, see Col. (7)).

In line with Bleakley & Lin (2012), we do not find that sunk investments matter and we argue that the rail lines helped solve the coordination problem of capital accumulation. In addition, we provide evidence that demographic growth reinforced these population effects. Yet should we conclude that railroads have initiated modern economic growth in Ghana? Per capita income has remained very low, around $1,350 (PPP, 2010 USD). Although railroad cells are more industrialized today (see Col. (6) of Table 5, e.g. +3.5*** for 0-10 km cells), Ghana’s employment share in manufacturing was only 6.4% in 2000. Manufacturing is relatively unproductive in Ghana, as it consists mostly of labor-intensive subsectors, such as African processed foods, furniture and clothing (Jedwab 2011, Jedwab & Osei 2012). That is why we characterize this effect as proto-industrialization. In other words, colonial railroads made some locations more attractive within Ghana but did not industrialize the country as a whole.
7. CONCLUSION

What is the impact of colonial infrastructure investments on long-term development? We investigate this issue by looking at the effects of railroad construction on development in Ghana. Using panel data at a fine spatial level over one century, we find a strong effect of rail connectivity on the production of cocoa, the country's main export commodity, and development, which we proxy by population and urban growth. First, we exploit various strategies to ensure our effects are causal. Second, transportation infrastructure investments had large welfare effects for Ghanaians during the colonial period. Colonization meant both extraction and development in this context. Third, rail construction had a persistent impact: railroad cells are more developed today despite a complete displacement of rail by other transportation means. Physical capital accumulation and the demographic transition account for path dependence.

We believe the paper makes the following contributions to the literature. First, we study colonial investments in physical capital instead of colonial institutions. Second, we look at transportation infrastructure investments, since they represent a very high share of colonial budgets. Third, we find strong evidence for path dependence and discuss how these colonial investments shaped the economic geography of Ghana. Fourth, we contribute to the literature on multiple spatial equilibria, by examining how the channels of path dependence could be different in a poor developing economy. Fifth, to our knowledge, we are the first paper to use population panel data at a fine spatial level over one century for an African country. This data set will help study why Africa is still underdeveloped.

Lastly, we believe that the patterns found for Ghana are paralleled in other African countries, e.g. coffee and tea in Kenya or groundnuts in Senegal. Colonial investments in transportation infrastructure shaped the economic geography of African countries, with consequences for long-run development.
REFERENCES


Jedwab, Remi, and Robert Darko Osei. 2012. “Structural Change in Ghana 1960-2010.” Case study prepared for the project “Structural Change in Developing Countries” under the supervision of Margaret McMillan and Dani Rodrik.


Figure 1: Cocoa Suitability, Provinces, and Railroad Lines in 1918.

Note: The suitable area corresponds to the tropical forest in 1900. Province boundaries date from 1916. Accra is the Gold Coast colony's capital. Kumasi is the hinterland capital. See Web Appendix 1 for data sources.
Figure 2: Railroad Lines in 1918 and Placebo Lines.

Note: The map only shows Southern Ghana. An 11x11 km cell is defined as suitable if it contains cocoa soils. The map displays the rail lines in 1918, the years line construction began and finished, the main rail stations and the placebo lines: lines that were planned but not built (Cape Coast-Kumasi 1873, Saltpond-Kumasi 1893, Apam-Kumasi 1897, Accra-Kumasi 1897, Accra-Kpong 1898) and lines that were not built early enough to affect cocoa production in 1927 (Tafo-Kumasi 1923, Huni Valley-Kade 1926). It also shows gold mines (Tarkwa, Obuasi and Kibi) and diamond fields (Kade). See Web Appendix 1 for data sources.
Figure 3: Transportation Networks in 1900 and Area of Profitable Production.

Note: The map shows transportation networks in 1900, the cells suitable for cocoa cultivation and the cells for which cocoa cultivation is profitable given transportation costs in 1900 and the producer price offered at the port in 1920-1929. Web Appendix 3 describes how this map was created. See Web Appendix 1 for data sources.

Figure 4: Cocoa Production, Exports and Transportation, 1900-1927.

Note: The figure displays three-year moving averages for cocoa production, cocoa tonnages transported by rail to a coastal port, and the share of cocoa exports of total exports from 1900 to 1927. See Web Appendix 1 for sources.
Figure 5: Railroad Lines in 1918, Cocoa Suitability, and Production in 1927.

Note: The map only shows Southern Ghana. A cell is defined as suitable if it contains cocoa soils and highly suitable if more than 50% of its area consists of forest ochrosols, the best cocoa soils. The map displays the railroad lines in 1918, suitable cells, highly suitable cells, and cocoa production in 1927. Each dot represents 100 tons of cocoa production. See Web Appendix 1 for data sources.
Figure 6: Railroad Lines in 1918 and Cell Population in 1901, 1931 and 2000.

Note: The maps displays cell population in 1901, 1931 and 2000. See Web Appendix 1 for data sources.
Figure 7: Long-Term Effects (Relative to 1901) of Rail Connectivity

Note: This graph displays estimates of Equation (4) for each distance threshold and each year = [1931, 1970, 2000], using 1901 as the reference year. We show the rail effects on urban population for additional years 1948, 1960 and 1984.
### TABLE 1: OBSERVABLES FOR TREATED CELLS VERSUS CONTROL CELLS

<table>
<thead>
<tr>
<th>RHS Variable:</th>
<th>Dummy Rail 1918, 0-20 km</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control Cells:</td>
<td>All</td>
</tr>
<tr>
<td>LHS Variable:</td>
<td>(1)</td>
</tr>
</tbody>
</table>

**Panel A: Economic Variables**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Column (1)</th>
<th>Column (2)</th>
<th>Column (3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mine Dummy</td>
<td>0.02**</td>
<td>0.01</td>
<td>0.02</td>
</tr>
<tr>
<td>Cocoa production in 1901</td>
<td>9.6**</td>
<td>9.6</td>
<td>9.6</td>
</tr>
</tbody>
</table>

**Panel B: Demographic Variables**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Column (1)</th>
<th>Column (2)</th>
<th>Column (3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Urban population in 1901</td>
<td>249</td>
<td>-276</td>
<td>231</td>
</tr>
<tr>
<td>Rural population in 1901</td>
<td>647***</td>
<td>379**</td>
<td>245</td>
</tr>
</tbody>
</table>

**Panel C: Physical Geography Variables**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Column (1)</th>
<th>Column (2)</th>
<th>Column (3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Share soils suitable for cocoa (%)</td>
<td>-0.11***</td>
<td>-0.14***</td>
<td>-0.14**</td>
</tr>
<tr>
<td>Share soils highly suitable for cocoa (%)</td>
<td>-0.21***</td>
<td>-0.26***</td>
<td>-0.32***</td>
</tr>
<tr>
<td>Share soils very highly suitable for cocoa (%)</td>
<td>-0.01</td>
<td>-0.04</td>
<td>0.01</td>
</tr>
<tr>
<td>Altitude: mean (m)</td>
<td>-18.0*</td>
<td>-1.3</td>
<td>48.0***</td>
</tr>
<tr>
<td>Altitude: standard deviation (m)</td>
<td>0.4</td>
<td>1.5</td>
<td>18.4***</td>
</tr>
<tr>
<td>Average annual rainfall (mm)</td>
<td>-22.4</td>
<td>21.6</td>
<td>77.6**</td>
</tr>
</tbody>
</table>

**Panel D: Economic Geography Variables**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Column (1)</th>
<th>Column (2)</th>
<th>Column (3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distance to Accra (km)</td>
<td>-42.1***</td>
<td>38.8***</td>
<td>31.6***</td>
</tr>
<tr>
<td>Distance to Kumasi (km)</td>
<td>-11.5**</td>
<td>2.1</td>
<td>-5.6</td>
</tr>
<tr>
<td>Distance to Aburi (km)</td>
<td>-39.4***</td>
<td>39.1***</td>
<td>18.8*</td>
</tr>
<tr>
<td>Distance to a port in 1901 (km)</td>
<td>-33.3***</td>
<td>5.0</td>
<td>29.2***</td>
</tr>
<tr>
<td>Distance to the coast (km)</td>
<td>-32.1***</td>
<td>5.9</td>
<td>28.6***</td>
</tr>
</tbody>
</table>

N Treated Cells: | 104 | 104 | 104 |
N Control Cells: | 450 | 152 | 44 |

**Notes:** Robust standard errors (not reported): * p<0.10, ** p<0.05, *** p<0.01. We regress each control variable on a dummy equal to one if the cell is less than 20 km from a 1918 railroad line. There are 15 different regressions for each column. In Column (1), all control cells are included. In Column (2), the control cells are the cells less than 20 km from a placebo line. In Column (3), the control cells are the cells less than 20 km from the Cape Coast-Kumasi placebo line (1873). See Web Appendix 1 for data sources.
TABLE 2: RAILROADS, COCOA PRODUCTION, AND POPULATION GROWTH

<table>
<thead>
<tr>
<th>Dependent Variable:</th>
<th>Cocoa 1927 (Tons)</th>
<th>Population 1931 (Number of Inhabitants)</th>
<th>Rural Population 1931 (in Localities &lt; 1,000 Inh.)</th>
<th>Urban Population 1931 (in Localities &gt; 1,000 Inh.)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
</tr>
<tr>
<td>Dummy Rail 1918, 0-10 km</td>
<td>666***</td>
<td>3,568***</td>
<td>2,405**</td>
<td>368</td>
</tr>
<tr>
<td></td>
<td>[114]</td>
<td>[1,112]</td>
<td>[1,056]</td>
<td>[650]</td>
</tr>
<tr>
<td>Dummy Rail 1918, 10-20 km</td>
<td>506***</td>
<td>1,385***</td>
<td>502</td>
<td>442</td>
</tr>
<tr>
<td></td>
<td>[110]</td>
<td>[513]</td>
<td>[507]</td>
<td>[510]</td>
</tr>
<tr>
<td>Dummy Rail 1918, 20-30 km</td>
<td>341***</td>
<td>782</td>
<td>187</td>
<td>139</td>
</tr>
<tr>
<td></td>
<td>[100]</td>
<td>[613]</td>
<td>[573]</td>
<td>[576]</td>
</tr>
<tr>
<td>Dummy Rail 1918, 30-40 km</td>
<td>246**</td>
<td>464</td>
<td>34</td>
<td>-4</td>
</tr>
<tr>
<td></td>
<td>[111]</td>
<td>[438]</td>
<td>[426]</td>
<td>[429]</td>
</tr>
<tr>
<td>Dummy Rail 1918, 40-50 km</td>
<td>55</td>
<td>203</td>
<td>107</td>
<td>67</td>
</tr>
<tr>
<td></td>
<td>[93]</td>
<td>[383]</td>
<td>[391]</td>
<td>[383]</td>
</tr>
<tr>
<td>Cocoa (Tons Produced) 1927</td>
<td>1.75***</td>
<td>1.77***</td>
<td>1.53***</td>
<td>1.13***</td>
</tr>
<tr>
<td>Dummy Rail Station 1918</td>
<td>4,106*</td>
<td>1,551</td>
<td>485</td>
<td>734</td>
</tr>
<tr>
<td></td>
<td>[2,102]</td>
<td>[1,746]</td>
<td>[561]</td>
<td>[591]</td>
</tr>
<tr>
<td>Cocoa at Rail St. (Tons) 1918</td>
<td>0.86*</td>
<td>-0.06</td>
<td>-0.06</td>
<td>0.91***</td>
</tr>
<tr>
<td>Cell FE and Year FE</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Controls</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.61</td>
<td>0.73</td>
<td>0.75</td>
<td>0.77</td>
</tr>
</tbody>
</table>

Notes: Robust standard errors clustered at the cell level are reported in parentheses; * p<0.10, ** p<0.05, *** p<0.01. All regressions include cell and year fixed effects, and cell controls interacted with year dummies: 1931 mine dummy, cocoa production in 1901, urban population in 1901, rural population in 1901, share (%) of soils suitable / highly suitable / very highly suitable for cocoa cultivation, mean and standard deviation (m) of altitude, average annual rainfall (mm), and Euclidean distances (km) to Accra, Kumasi, Aburi, a port in 1901, and the coast. See Web Appendix 1 for data sources.
TABLE 3: ROBUSTNESS CHECKS

<table>
<thead>
<tr>
<th>Regression:</th>
<th>Main</th>
<th>Highly Suitable Cells</th>
<th>West vs. East</th>
<th>District-Year FE</th>
<th>Matching Cells</th>
<th>C. Coast</th>
<th>IV Straight Lines</th>
<th>Distance To Rail Station</th>
<th>No Controls</th>
<th>No Nodes</th>
<th>Full Sample</th>
<th>Logs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
<td>(6)</td>
<td>(7)</td>
<td>(8)</td>
<td>(9)</td>
<td>(10)</td>
<td>(11)</td>
<td>(12)</td>
</tr>
<tr>
<td>Panel A: Dependent Variable = Cocoa (Tons Produced) in 1927</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dummy Rail 1918, 0-20 km</td>
<td>556*** 929*** 316*** 747*** 419*** 775*** 731*** 622*** 571*** 544*** 552*** 424*** 2.06***</td>
<td>[85] [168] [79] [224] [104] [121] [155] [170] [88] [113] [87] [77] [0.40]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dummy Rail 1918, 20-40 km</td>
<td>275*** 391*** 115 326*** 221*** 407*** 421*** 35 260*** 373*** 275*** 194*** 0.96**</td>
<td>[78] [111] [71] [122] [83] [106] [149] [313] [80] [94] [78] [66] [0.41]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rail 1918 x East. Line, 0-20 km</td>
<td>781*** 334</td>
<td>[227] [308]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rail 1918 x East. Line, 20-40 km</td>
<td>509** 171</td>
<td>[200] [272]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Panel B: Dependent Variable = Population (Number of Inhabitants) in 1931

| Dummy Rail 1918, 0-20 km | 2,034***2,387***1491* 1950* 1,496*** 2,052*** 1,754*** 2,944*** 2,027*** 1,939*** 1,486*** - 0.71** | [605] [858] [624] [1059] [561] [714] [720] [1,328] [622] [563] [427] [0.35] |
| Rail 1918 x East. Line, 0-20 km | 1,943 815 | [1,434] [1,749] |

| Cell FE and Year FE | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y |
| Controls | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y |

Notes: Robust standard errors clustered at the cell level are reported in parentheses; * p<0.10, ** p<0.05, *** p<0.01. All regressions include cell and year fixed effects and cell controls interacted with year dummies. These are the same as in Table 2. In Column (2), the sample is restricted to highly suitable cells. In Columns (3) and (4), we compare the effects for the Western and Eastern Lines. In Column (4), we restrict the sample to highly suitable cells when comparing both lines. In Column (5), we include district-year fixed effects. In Columns (6) and (7), the sample of control cells is restricted to all placebo cells and cells along the Cape-Coast placebo line (1873) respectively. In Column (8), we instrument the rail dummies by the distance to the straight lines Sekondi-Kumasi and Accra-Kumasi. In Column (9), the rail dummies are generated using the distance to rail stations. In Column (10), we drop the controls. In Column (11), we drop the railroad nodes. In Column (12), we use the full sample, but the data set is not exhaustive for total population. In Columns (13), we use a log-linear functional form. See Web Appendix 1 for data sources.
### TABLE 4: PLACEBO LINES

<table>
<thead>
<tr>
<th>Type of Placebo Line:</th>
<th>Planned But Never Built (From West to East)</th>
<th>Not Built Yet</th>
<th>All Lines</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>C.Coast Kумasi 1873</td>
<td>Saltpond Kумasi 1893</td>
<td>Apam Kумasi 1897</td>
</tr>
<tr>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
</tr>
</tbody>
</table>

#### Panel A: Dependent Variable = Cocoa (Tons Produced) in 1927


#### Panel B: Dependent Variable = Cocoa (Tons Produced) in 1927


#### Panel C: Dependent Variable = Population (Number of Inhabitants) in 1931


#### Panel D: Dependent Variable = Population (Number of Inhabitants) in 1931

| Controls | Y | Y | Y | Y | Y | Y | Y |

Notes: Robust standard errors clustered at the cell level are reported in parentheses; * p<0.10, ** p<0.05, *** p<0.01. All regressions include cell and year fixed effects and cell controls interacted with year dummies. These are the same as in Table 2. In Panels A and D, we compare the placebo cells with the non-placebo cells (the railroad cells and the other control cells). In Panels B and E, we compare the placebo cells that do not intersect with a 1918 line with the other cells. In Panels C and F, we drop the cells less than 20 km from the railway lines in 1918, in order to compare the placebo cells with the other control cells. See Web Appendix 1 for data sources.
### TABLE 5: RAILROADS AND ECONOMIC DEVELOPMENT IN 2000

<table>
<thead>
<tr>
<th>Dependent Variable:</th>
<th>Urban Population (Inh.)</th>
<th>Lights (%)</th>
<th>Solid Walls (%)</th>
<th>Employment Share (%)</th>
<th>Post Office (%)</th>
<th>Telephone (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No nodes</td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
</tr>
<tr>
<td>Dummy Rail 1918, 0-10 km</td>
<td>33,791*</td>
<td>10,059***</td>
<td>29.8***</td>
<td>14.1***</td>
<td>-16.0***</td>
<td>3.5***</td>
</tr>
<tr>
<td></td>
<td>[17,862]</td>
<td>[3,844]</td>
<td>[4.7]</td>
<td>[3.0]</td>
<td>[2.8]</td>
<td>[0.8]</td>
</tr>
<tr>
<td>Dummy Rail 1918, 10-20 km</td>
<td>3,785</td>
<td>3,231**</td>
<td>16.0***</td>
<td>4.6*</td>
<td>-7.7***</td>
<td>2.1***</td>
</tr>
<tr>
<td></td>
<td>[2,767]</td>
<td>[1,373]</td>
<td>[4.3]</td>
<td>[2.4]</td>
<td>[2.3]</td>
<td>[0.7]</td>
</tr>
<tr>
<td>Dummy Rail 1918, 20-30 km</td>
<td>1,441</td>
<td>1,361</td>
<td>7.4**</td>
<td>1</td>
<td>-5.4***</td>
<td>1.7***</td>
</tr>
<tr>
<td></td>
<td>[2,719]</td>
<td>[1,501]</td>
<td>[3.6]</td>
<td>[2.1]</td>
<td>[2.0]</td>
<td>[0.6]</td>
</tr>
<tr>
<td>Dummy Rail 1918, 30-40 km</td>
<td>1,046</td>
<td>858</td>
<td>3.7</td>
<td>2.7</td>
<td>-2.7</td>
<td>0.9</td>
</tr>
<tr>
<td></td>
<td>[2,222]</td>
<td>[1,281]</td>
<td>[3.2]</td>
<td>[2.2]</td>
<td>[2.0]</td>
<td>[0.6]</td>
</tr>
<tr>
<td>Dummy Rail 1918, 40-50 km</td>
<td>4,633</td>
<td>1,569</td>
<td>0</td>
<td>4.8*</td>
<td>-4.3**</td>
<td>1.6***</td>
</tr>
<tr>
<td></td>
<td>[3,143]</td>
<td>[1,406]</td>
<td>[3.0]</td>
<td>[2.7]</td>
<td>[2.0]</td>
<td>[0.6]</td>
</tr>
<tr>
<td>Mean</td>
<td>10,551</td>
<td>8,235</td>
<td>29.4</td>
<td>19.6</td>
<td>74.8</td>
<td>6.4</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Dependent Variable:</th>
<th>Literate (%)</th>
<th>SSS Educ. (%)</th>
<th>Primary School (%)</th>
<th>JSS (%)</th>
<th>SSS (%)</th>
<th>Clean Water (%)</th>
<th>Clinic (%)</th>
<th>Hospital (%)</th>
<th>Paved Road (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dummy Rail 1918, 0-10 km</td>
<td>7.6***</td>
<td>2.9**</td>
<td>7.1*</td>
<td>10.8***</td>
<td>12.3**</td>
<td>7.8**</td>
<td>9.1**</td>
<td>19.9***</td>
<td>0.31***</td>
</tr>
<tr>
<td></td>
<td>[1.9]</td>
<td>[1.2]</td>
<td>[3.6]</td>
<td>[3.7]</td>
<td>[5.0]</td>
<td>[3.6]</td>
<td>[4.6]</td>
<td>[4.8]</td>
<td>[0.08]</td>
</tr>
<tr>
<td>Dummy Rail 1918, 10-20 km</td>
<td>2.4</td>
<td>-0.1</td>
<td>7.9**</td>
<td>9.6***</td>
<td>8.9**</td>
<td>0.2</td>
<td>5.4</td>
<td>10.7***</td>
<td>0.23***</td>
</tr>
<tr>
<td></td>
<td>[1.6]</td>
<td>[1.0]</td>
<td>[3.1]</td>
<td>[3.2]</td>
<td>[3.6]</td>
<td>[2.5]</td>
<td>[4.0]</td>
<td>[3.5]</td>
<td>[0.07]</td>
</tr>
<tr>
<td>Dummy Rail 1918, 20-30 km</td>
<td>1.3</td>
<td>0.1</td>
<td>4.7</td>
<td>6.6**</td>
<td>1.7</td>
<td>1.1</td>
<td>0.4</td>
<td>7.2**</td>
<td>0.21***</td>
</tr>
<tr>
<td></td>
<td>[1.6]</td>
<td>[0.9]</td>
<td>[2.9]</td>
<td>[3.2]</td>
<td>[3.7]</td>
<td>[2.7]</td>
<td>[4.0]</td>
<td>[3.3]</td>
<td>[0.07]</td>
</tr>
<tr>
<td>Dummy Rail 1918, 30-40 km</td>
<td>0.5</td>
<td>-0.6</td>
<td>5.1*</td>
<td>5.8*</td>
<td>-0.5</td>
<td>2.1</td>
<td>-0.9</td>
<td>5.3*</td>
<td>0.14**</td>
</tr>
<tr>
<td></td>
<td>[1.5]</td>
<td>[0.8]</td>
<td>[2.7]</td>
<td>[3.1]</td>
<td>[3.5]</td>
<td>[2.3]</td>
<td>[4.0]</td>
<td>[3.1]</td>
<td>[0.06]</td>
</tr>
<tr>
<td>Dummy Rail 1918, 40-50 km</td>
<td>1.2</td>
<td>-0.4</td>
<td>5.1*</td>
<td>6.0*</td>
<td>-0.6</td>
<td>4.0</td>
<td>1.0</td>
<td>4.7</td>
<td>0.16**</td>
</tr>
<tr>
<td></td>
<td>[1.5]</td>
<td>[0.9]</td>
<td>[2.8]</td>
<td>[3.2]</td>
<td>[3.7]</td>
<td>[2.5]</td>
<td>[4.1]</td>
<td>[3.3]</td>
<td>[0.06]</td>
</tr>
<tr>
<td>Mean</td>
<td>52.0</td>
<td>13.4</td>
<td>84.0</td>
<td>74.5</td>
<td>22.5</td>
<td>14.2</td>
<td>43.4</td>
<td>14.6</td>
<td>0.29</td>
</tr>
</tbody>
</table>

Notes: Robust standard errors are reported in parentheses; * p<0.10, ** p<0.05, *** p<0.01. We use various sources to reconstruct 18 development outcomes for 554 cells around 2000. All regressions include the same controls as in Table 2. Col. (1) and (2): urban population (number of inhabitants), with and without the railroad nodes. Col. (3): share of cell area (%) for which a light is observed by satellite. Col. (4): share of inh. (%) in a residence with solid walls. Col. (5)-(7): sectoral employment shares (%). Col. (8), (9), (12), (13), (14), (16) and (17): share of inh. (%) living less than 5 km from a: post office, telephone line, primary school, junior secondary school (JSS), senior secondary school (SSS), health clinic or hospital. Col. (10)-(11): shares of adults (≥ 18 y.o.) that are literate, have attended a SSS. Col. (15): share of inh. (%) with access to clean water. Col. (16): dummy if the cell is crossed by a paved road. See Web Appendix 1 for data sources.
<table>
<thead>
<tr>
<th>Dependent Variable:</th>
<th>Log Urban Population (Number of Inh. in Loc. &gt; 1,000 Inh.)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Localities: 5,000- + 5,000- + 1,000- 1,000- 1,000-</td>
</tr>
<tr>
<td></td>
<td>(1) (2) (3) (4) (5) (6) (7) (8) (9)</td>
</tr>
<tr>
<td>Dummy Rail 1918, 0-10 km</td>
<td>1.21*** 0.13 0.22 0.51 0.07 -0.93** -0.32 -0.08 -0.12</td>
</tr>
<tr>
<td></td>
<td>[0.45] [0.41] [0.44] [0.34] [0.33] [0.45] [0.41] [0.45] [0.41]</td>
</tr>
<tr>
<td>Dummy Rail 1918, 10-20 km</td>
<td>0.91* 0.25 0.15 0.45 0.30 0.05 0.35 0.39 0.28</td>
</tr>
<tr>
<td></td>
<td>[0.49] [0.44] [0.44] [0.37] [0.35] [0.36] [0.34] [0.38] [0.37]</td>
</tr>
<tr>
<td>Dummy Rail 1918, 20-30 km</td>
<td>0.24 -0.40 -0.49 -0.45 -0.54 -0.01 0.15 -0.45 -0.47</td>
</tr>
<tr>
<td></td>
<td>[0.54] [0.47] [0.48] [0.33] [0.33] [0.39] [0.35] [0.37] [0.36]</td>
</tr>
<tr>
<td>Dummy Rail 1918, 30-40 km</td>
<td>0.26 0.11 0.30 0.24 0.18 -0.35 -0.16 0.39 0.31</td>
</tr>
<tr>
<td></td>
<td>[0.46] [0.44] [0.45] [0.33] [0.31] [0.32] [0.29] [0.35] [0.33]</td>
</tr>
<tr>
<td>Dummy Rail 1918, 40-50 km</td>
<td>0.62 0.44 0.57 0.31 0.13 0.49 0.45 -0.01 0.04</td>
</tr>
<tr>
<td></td>
<td>[0.45] [0.39] [0.39] [0.30] [0.28] [0.41] [0.35] [0.35] [0.34]</td>
</tr>
<tr>
<td>Log Urban Population, 1901</td>
<td>0.21*** 0.08* 0.03 0.03 0.00 0.06 -0.04 0.04 0.04</td>
</tr>
<tr>
<td></td>
<td>[0.04] [0.04] [0.04] [0.03] [0.03] [0.08] [0.06] [0.05] [0.04]</td>
</tr>
<tr>
<td>Log Rural Population, 1901</td>
<td>0.62*** 0.19*** 0.16** 0.06 0.01 0.07 0.08* -0.01 0.01</td>
</tr>
<tr>
<td></td>
<td>[0.06] [0.06] [0.06] [0.05] [0.04] [0.04] [0.04] [0.05] [0.05]</td>
</tr>
<tr>
<td>Log Urban Population, 1931</td>
<td>0.22*** 0.19*** -0.00 -0.01 0.17*** 0.10* -0.06 -0.04</td>
</tr>
<tr>
<td></td>
<td>[0.04] [0.04] [0.02] [0.02] [0.05] [0.05] [0.05] [0.03] [0.03]</td>
</tr>
<tr>
<td>Log Rural Population, 1931</td>
<td>0.60*** 0.58*** 0.21*** 0.16*** -0.05 -0.03 0.24*** 0.21***</td>
</tr>
<tr>
<td></td>
<td>[0.07] [0.06] [0.06] [0.05] [0.03] [0.03] [0.06] [0.06]</td>
</tr>
<tr>
<td>Log Urban Population, 1970</td>
<td>0.08** 0.04 0.06 0.06* 0.03 0.02</td>
</tr>
<tr>
<td></td>
<td>[0.03] [0.03] [0.04] [0.03] [0.03] [0.04] [0.04]</td>
</tr>
<tr>
<td>Log Rural Population, 1970</td>
<td>0.26** 0.10 -0.06 -0.02 0.08 0.05</td>
</tr>
<tr>
<td></td>
<td>[0.11] [0.12] [0.09] [0.06] [0.14] [0.14]</td>
</tr>
<tr>
<td>Log Urban Population, 1984</td>
<td>0.51*** 0.42*** 0.03 0.39***</td>
</tr>
<tr>
<td></td>
<td>[0.05] [0.05] [0.03] [0.05]</td>
</tr>
<tr>
<td>Log Rural Population, 1984</td>
<td>0.35** 0.33** 0.13 0.09 0.46** 0.48**</td>
</tr>
<tr>
<td></td>
<td>[0.14] [0.14] [0.11] [0.10] [0.19] [0.19]</td>
</tr>
<tr>
<td>Mfg. Employment Share (%)</td>
<td>0.05* 0.12*** 0.08** -0.01 0.02</td>
</tr>
<tr>
<td></td>
<td>[0.03] [0.03] [0.03] [0.05] [0.04]</td>
</tr>
<tr>
<td>Serv. Employment Share (%)</td>
<td>0.03* 0.08*** 0.07*** 0.04** 0.03*</td>
</tr>
<tr>
<td></td>
<td>[0.01] [0.01] [0.01] [0.02] [0.01]</td>
</tr>
<tr>
<td>Log Urban Pop. 1,000-5,000, 1984</td>
<td>0.03 0.42***</td>
</tr>
<tr>
<td></td>
<td>[0.02] [0.05]</td>
</tr>
<tr>
<td>Log Urban Pop. 5,000--+, 1984</td>
<td>0.53*** -0.00</td>
</tr>
<tr>
<td></td>
<td>[0.05] [0.04]</td>
</tr>
</tbody>
</table>

| Sunk investments 1901-1931 | - - Y Y Y Y Y Y Y |
| Sunk investments 2000 | - - - - Y Y Y Y |
| Human capital 2000 | - - - - Y Y Y Y |
| Observations | 552 552 552 552 552 552 552 552 552 |
| R-squared | 0.34 0.50 0.50 0.71 0.74 0.67 0.74 0.64 0.67 |

**Notes**: Robust standard errors are reported in parentheses; * p < 0.10, ** p < 0.05, *** p < 0.01. Sunk investments 1901-1931: number of government and non-government schools, European and African hospitals and dummy for being crossed by a class 1 or class 2 road in 1901 and 1931. Sunk investments 2000: share of inh. (%) in a residence with solid walls, share of inh. (%) living less than 5 km from a post office, telephone line, primary school, junior secondary school, senior secondary school (SSS), health clinic or hospital, share of inh. (%) with access to clean water, and dummy if the cell is crossed by a paved road. Human capital 2000: shares of adults that are literate, have attended a SSS. See Web Appendix 1 for data sources.
Web Appendix 1: Data Description

This appendix describes in detail the data we use in our analysis.

Spatial Units:
We assemble data for 2,091 grid cells of 0.1x0.1 degrees (11x11 km) in Ghana from 1891 to 2000. We choose a high resolution grid because we have very precise GIS data on railways, cocoa production, population, and urban growth in 1891-2000, and economic development in 2000. Each grid cell has the same size, except those cells that are coastal or crossed by a border. We create two dummies equal to one if the grid cell is coastal or bordering another country to control for this issue. Grid cells belong to 110 districts and 10 provinces using 2000 boundaries, or 4 regions using 1931 boundaries.

Railway Data:
We obtain the layout of railway lines in GIS from Digital Chart of the World. We then use Gould (1960), Dickson (1968), Tsey (1986), and Luntinen (1996) to recreate the history of railway construction. For each line, we know when it was surveyed, planned, started, and finished and when each station was reached and opened. From the same sources, we know lines that were built but not planned. Most of those placebo lines follow historical trade routes and became roads later. Using the GIS road network also available from Digital Chart of the World, we recreate those placebo lines in GIS. We calculate for each grid cell the Euclidean distance (km) from the cell centroid to each real or placebo line. Lastly, we create a set of dummies equal to one if the grid cell is less than X km away from the railway line: 0-10, 10-20, 20-30, 30-40 and 40-50 km. We create a dummy equal to one if the grid cell contains a rail station in 1918. We also know how many tons of cocoa were brought to each station in 1918.\(^1\) Data on railway traffic for both lines was obtained from various sources.\(^2\) Lastly, we construct the instruments by using GIS to obtain the Euclidean distance (km) from each cell centroid to the straight lines Sekondi-Kumasi, Accra-Sekondi, Sekondi-Tarkwa-Obuasi and Accra-Kibi.

Cash Crop Production and Price Data:
A very precise map of cocoa production in 1927 was obtained from the 1927 Yearbook of the Gold Coast and digitized. This map displays dots for each 100 tons of cocoa production.\(^3\) We then use GIS to reconstruct total cocoa production (tons) for each grid cell in 1927. We proceed similarly to create cocoa production in 1950 using a production map published in the Report on the Cocoa Industry in Sierra Leone and Notes on the Cocoa Industry of the Gold Coast by Cadbury Brothers LTD in 1955. We use Bateman (1965) to obtain the international and producer prices.

We collect population data from the gazetteers of the Population and Housing Censuses 1891, 1901, 1931, 1948, 1960, 1970, 1984 and 2000. They list localities and their population size. Defining as a city any locality with more than 1,000 inhabitants, we obtain a geospatialized sample of 1,373 different cities in Ghana for all these years. Using GIS, we recalculate total urban population for each grid cell. We are then able to recreate rural population for each gridcell in 1901, 1931, 1970, 1984 and 2000. From the census gazetteer, we know the population size of each village (locality with less than 1,000 inhabitants). It was impossible to find the geographical coordinates of all of them. Yet, the 1901 census was exhaustively conducted and geospatialized in the South of Ghana (756 cells). We know for each cell the number of large towns, towns (more than 500 inhabitants), head chief towns, large villages (100-500 inhabitants) and villages (less than 100 inhabitants). Using GIS, we can deduce for each cell the number of villages that are less than 100 inhabitants, the number of villages that have between 100 and 500 inhabitants and the number of villages that have between 500 and 1,000 inhabitants. From the census, we know the average settlement size for each category and we can reconstruct total rural population for each cell in 1901. For 1931, we have a map of the distribution of population for the whole country.\(^4\) This map displays at a very fine spatial level settlements that have less than 500 inhabitants and settlements that have between 500 and 1,000 inhabitants. From the census, we know the average settlement size for each category, and we can reconstruct total rural population for each cell in 1931. We use the 2000 Facility Census which has population data on all settlements in 1970, 1984 and

---

\(^1\)This information was retrieved from the Administration Report of Gold Coast Railways for the Year 1921.
\(^3\)Aggregating all the dots, we obtain 209,100 tons of cocoa production in 1927, which is very comparable with what we find from national estimates (see Gunnarsson 1978).
\(^4\)The map was obtained from the 1960 Ghana Population Atlas.
Demographic Data:
Access to clean water. As described above, we also have data on roads in 1901, 1931 and 2000. Junior secondary school (JSS), senior secondary school (SSS), health clinic or hospital, and share of inhabitants with grid cell level in 2000: share of inhabitants living less than 5 km from a post office, telephone line, primary school, 2000
Facility Census
Gold Coast for the years 1902 and 1929-30. Besides, we use the Colonial Annual Reports of the Gold Coast, published in 1930 by the Survey Headquarters. Those road map have a consistent legend showing class 1 roads ("roads suitable for motor traffic throughout the year"), class 2 roads ("roads suitable for motor traffic but occasionally closed"), and class 3 roads ("roads suitable for motor traffic in dry season only"). Other roads are not suitable for motor traffic and are not considered here. We use a Michelin paper map to recreate the 2000 road network in GIS. We also have data on European and African hospitals in 1902 and 1930. The data was compiled from health reports:

Infrastructure Data in 1901, 1931 and 2000:
We use annual production data for Ghanaian mines in 1901-1931 for three commodities: gold, manganese, and diamonds. As we have the geographical coordinates of each mine, we create a dummy equal to one if the cell contains a mine.

Geographical Data:
Forest data comes from land cover GIS data compiled by Globcover (2009). The data displays those areas with virgin forest or mixed virgin forest/croplands, which were areas with virgin forest before it was cleared for cash crop production. Soil data comes from the 1958 Survey of Ghana Classification Map of Cocoa Soils for Southern Ghana. This map was digitized in GIS and we calculated for each cell the share of land which is suitable for cocoa cultivation. We also know the respective shares of land which consists of ochrosols (first class, second class, third class, unsuitable), oxysols, and intergrades. A cell is defined as suitable if it contains cocoa soils. It is then highly suitable if more than 50% of its area consists of forest ochrosols, the best soils for cocoa cultivation. It is very highly suitable if more than 50% of its area consists of class 1 and class 2 ochrosols. Climate data comes from Terrestrial Air Temperature and Precipitation: 1900-2007 Gridded Monthly Time Series, Version 1.01, 2007, University of Delaware. We estimate for each grid cell average annual precipitations (mms) in 1900-1960. Topography comes from SRTM3 data. We estimate for each grid cell the mean and standard deviation of altitude (meters). The standard deviation captures the slope and ruggedness of the terrain.

Economic Geography Data:
For each grid cell, we use GIS to get the Euclidean distances (km) to Accra, the capital city, Kumasi, the largest hinterland city, Aburi, the city where Ghanaian cocoa production originated, a port in 1901, and the coast.

Transportation Networks in 1901 and Road Data in 1901-2000:
Transportation networks in 1901 are obtained from Gould (1960) and Dickson (1968), as well as the Colonial Annual Reports of the Gold Coast for the year 1903, 1904 and 1907. We use various sources to reconstruct a GIS database of roads in 1901, 1922 and 1931: Gould (1960) and Map of The Gold Coast with Togoland Under British Mandate, published in 1930 by the Survey Headquarters. Those road map have a consistent legend showing class 1 roads ("roads suitable for motor traffic throughout the year"), class 2 roads ("roads suitable for motor traffic but occasionally closed"), and class 3 roads ("roads suitable for motor traffic in dry season only"). Other roads are not suitable for motor traffic and are not considered here. We use a Michelin paper map to recreate the 2000 road network in GIS, distinguishing paved (bitumenized), improved (laterite), and earthen roads.

Infrastructure Data in 1901, 1931 and 2000:
We have data on government and non-government schools (missions) in 1902 (which we use as an approximation for 1901) and 1930-31. The data was compiled from education reports: Report on the Education Department of the Gold Coast for the years 1902 and 1930-31. They list all the schools in the country, which we then geospatialized. We also have data on European and African hospitals in 1902 and 1930. The data was compiled from health reports: Report on the Medical and Sanitary Department of the Gold Coast for the years 1902 and 1929-30. Besides, we use the 2000 Facility Census and the 2000 Population and Housing Census to recreate data on infrastructure provision at the grid cell level in 2000: share of inhabitants living less than 5 km from a post office, telephone line, primary school, junior secondary school (JSS), senior secondary school (SSS), health clinic or hospital, and share of inhabitants with access to clean water. As described above, we also have data on roads in 1901, 1931 and 2000.

Demographic Data:
Using various sources, we track the evolution of birth and death rates separately for urban and rural Ghana in 1901, 1911, 1921 and 1984. The main sources are: CICRED - The Population of Ghana 1974, Patterson (1979, 1981), Mining production and price data is collected from the following documents: The Mineral Industry of the British Empire and Foreign Countries 1913-1919; Reports of the Mines Department of the Gold Coast 1931-1958.

Height Data:
Data on Ghanaian soldiers enlisted in the British Army in 1888-1960 was collected by Alexander Moradi using official soldier files. The data set is described in Moradi (2009). For each soldier, we know his year and place of birth, the year he enlisted, his age, farming specialization (2 categories), the skill content of his previous occupation, (5 categories), his literacy skill level (2 categories), ethnicity (68 groups), and height.

Employment, Human Capital and Economic Development Data in 1931 and 2000:
Data on urban employment in 1931 comes from the Population and Housing Censuses. We use the 2000 Facility Census and the 2000 Population and Housing Census to recreate data on economic development at the grid cell level in 2000: sectoral employment shares (agriculture, industry/manufacturing, and services), share of literate adults (≥ 18 y.o.), having attended a SSS, share of inhabitants in a residence with solid walls, and share of cell area (%) for which a light is observed by satellite. The source of the satellite data on night lights is Henderson, Storeygard and Weil (2012).

Data on Railroads in Sub-Saharan Africa:
We obtain the layout of Sub-Saharan African rail lines in GIS from Digital Chart of the World. We then use Wikipedia and studies available on the internet to recreate the history of each line. We do not reproduce this list here as it would be too long. We know when each line was built, whether it was before the country became independent, and the main motivations behind its construction. There are 66,491 km of railway lines in 2000, but 58,716 km were built before independence (57,722 km if we use the year 1960 instead). There are three groups of motivation: military domination (against natives or other colonial powers), mining, and cash crop agriculture. The same line can serve several purposes. Data on colonial budgets in Kenya and French West Africa was compiled by Burgess et al. (2011) and Huillery (2010).

REFERENCES
Web Appendix 2: Placebo Lines

This appendix presents background information on the placebo lines. It draws heavily from Gould (1960), Tsey (1986) and Luntinen (1996). Figure 1 shows their location.

1 Lines Proposed But Never Built

Several private initiatives submitted proposals to convince the Colonial Office of the profitability of their rail schemes. All asked for a government guarantee of interest on capital outlay. The government opposed this, with reference to the obvious incentive problems of such a guarantee, subsidizing British investors and leading, as in India, to over-capitalization. These initiatives did not enter a phase of concrete planning, nor were they able to raise the capital necessary for their schemes. Yet pressure mounted to build railroads in Ghana, as the French and the Germans were already building their own networks in West Africa. Tsey (1986, p.17) writes: "Indeed this had become all the more urgent in view of the fact that the Brussels Conference of 1889-1890 had called on all the European powers to back up their claims to colonial territories in Africa by the establishment of effective administrations particularly through the construction of railways and roads." Eventually, the state favored public ownership. The government proposals described below can be considered as counterfactual alternatives to the two routes built. We argue that these lines all had the same probability of being built, and only random events (e.g., a change in the colonial Governor) explain why construction did not happen.

Cape Coast-Kumasi 1873: The first proposal to build a railroad was made in 1873 to connect Cape Coast to Kumasi via Prasu. The British planned to use the line to send troops to fight the Ashanti. Some rail materials were landed at Cape Coast, but the project was dropped, since it was not possible to build the line in time for the military operation. The line was also proposed in 1891 and 1892 by private consortia (Luntinen 1996, p.18). Cape Coast was the capital of the Gold Coast colony before it was moved to Accra in 1877. The latter had a drier climate and was believed to be a healthier place for Europeans. In 1901 Cape Coast was similar in size to Accra (respectively 28,948 and 30,144). Moreover, Cape Coast was the starting point of an important historical trade route to Kumasi. Villages clustered along this road like pearls on a string. This permitted Cape Coast to achieve the largest trade volume of the coastal towns in 1900 (Gould 1960, p.17). Hence, in terms of existing traffic, the line also had some potential.

Saltpond-Oda-Kumasi 1893: In 1893, the colonial government commissioned a survey for a railroad network that would benefit the whole country, not only the mining industry. Government and engineers favored a trunk route. Saltpond was chosen as starting point for its central location and because construction materials could be landed easily. The line was to reach Kumasi, crossing densely populated and rich palm kernel and palm oil areas. The line had the support of Governor Griffith. Tsey (1986, p.19) writes: "A major step towards a full railway proposal came when Governor Griffith suggested a 300-400 mile network, the first phase of which was to form a trunk route from the most central port of the colony into the interior. Such a central line through the agricultural districts, the Governor argued, not only offered some hope of financial return, but in addition, it would form a suitable basis for future railway development." Yet Governor Griffith retired in April 1895, and the new Governor Maxwell changed course and again favored the mining industry. Tsey (1986, p.32) explains that Governor Maxwell, who previously worked in the Malay States where railroads served the tin mines, thought that the same model should be applied to Ghana with gold.

Apam-Oda-Kumasi 1897, Accra-Oda-Kumasi 1897: Although the construction of the Western line to mining areas was about to start, it was still widely accepted that the colony needed a central line to be built to tap the palm oil areas. Yet the Saltpond-Oda-Kumasi project was eventually dropped because of the relatively higher capital outlay compared to Apam as starting point. The Apam-Oda-Kumasi line was slightly shorter, thus cheaper, and had all the advantages of a central rail route. Calculations of the consulting engineer indicated its profitability. However, discontent grew amongst Accra merchants who thought that the capital and chief trading town should be connected first. The Accra-Oda-Kumasi was also surveyed, and "the railway surveys led to the conclusion that Apam was the better terminus for a Kumasi railway" (Luntinen 1996, p.33). A conference was to be held in London before the end of 1897 to discuss the various proposals but unexpectedly Governor Maxwell died before reaching London. He was replaced by Governor Hodgson who "was of the opinion that the principal railways of the colony should converge upon Accra. This town had been selected as the seat of the Government, and it was the most healthy site on the coast for Europeans. Much money had been invested there and he thought that it would be a mistake to raise a rival port." (Luntinen 1996, p.34). Besides, he thought that a central line from Accra would be useless as it would not directly traverse the palm oil areas below Apam, so he favored another line to Kpong, on the Volta River. The possible extension of the Sekondi-Tarkwa line - though originally intended as a short, local mining railroad - to
Kumasi also undermined the central route strategy.

**Accra-Kpong 1898:** The main objective of this line was to tap palm oil areas to the north-east of Accra and boost cotton cultivation in the districts bordering Togo. It would also link Accra to the Volta river, to facilitate the transport of government stores to the North, and to the “Government Sanatorium and Botanical Station at Aburi, 1400 feet above sea level, [...] so that Europeans will be able to reside in this delightful spot, coming daily to their offices in Accra” (Tsey 1986, p.55). The line was approved for construction, but it was decided that works would not begin before the extension to Kumasi was completed in late 1903. However, Governor Nathan decided to retire in February 1904. He was replaced by Governor Rodger, who thought that a rail line from Accra to Kumasi was more important and that the Volta river was already being used from the coastal port of Apam. The mercantile representative on the Legislative Council also strongly opposed the line, given “he was operating a boat service on the Volta” between Kpong and Apam and did not want the competition of the rail (Tsey 1986, p.57).

### 2 Lines Not Yet Built

The rail network was later expanded. We also consider lines that were actually built, but not in time to affect production in 1927. Cocoa is a perennial crop. Pod production starts after 5 years and peaks after 25 years. Hence, to observe an impact on cocoa production in 1927, farmers must have grown cocoa trees before 1923. There is no qualitative evidence that this happened to a significant degree. If the prospect of rail connectivity did indeed induce much production in advance, it would add to the positive correlation expected from reverse causality (cocoa production attracted the railroad). However, we do not find any positive effect of the two placebo lines below.

**Tafo-Kumasi 1923:** The Eastern line, with Accra as terminus, reached Tafo in 1916, when war time restrictions on construction materials suspended all further rail projects. Tafo station was opened in July 1917. Bauxite discoveries, midway between Tafo and Kumasi in 1917, led to the decision to extend the railroad to Kumasi (Tsey 1986, p.64). Actual construction, however, only started in 1920. The line was completed in 1923.

**Huni Valley-Kade 1927:** The line ran parallel to the coast, about 80 km inland. It connected the diamond mines at Kade and was supposed to encourage cocoa, kola, palm oil, and timber exports. By conveying more traffic to the newly developed harbor at Takoradi (Sekondi), it hoped to make the port viable. Construction begun in 1923. Several roads already connected the area to the coast, but they were of poor quality. Railroad surveyors believed that lorry traffic could not operate profitably beyond 50 km from the coast, but this turned out to be wrong. The short distance to the coast made lorry transport very competitive reviving the old ports of Cape Coast, Saltpond and Winneba directly to the South.

### REFERENCES


Web Appendix 3: Cost Maps

Using the Path Distance function in ArcGIS we computed the lowest cost path from every location to coastal ports. In this exercise, locations correspond to grid cells of 90m x 90m. This resolution was determined by the largest cell size in our analysis which was the elevation raster (Jarvis, Reuter, Nelson and Guevara 2008). We then averaged the costs derived at the 90mx90m resolution to the corresponding 0.1x0.1 degree grids that we used in the main analysis. All cocoa was produced for export. Hence, coastal ports were the terminal points of the produce. In 1900 there was little concentration; commodities were exported from 19 coastal towns, as shown in Figure A3.1 (Gould 1960, p. 45).

**Cost layer:** The cost layer assigns costs of movement (transport) through one cell relative to other cells. Transportation costs varied not only by transportation method, but also by bulkiness, labour costs, and terrain. Wages of carriers, for example, were frequently cited as 1.3s per day in the Gold Coast whereas the same carrier would earn as little as 0.75s in the Northern Territories, e.g. from Kintampo to Gambaga (Gold Coast 1907; Gould 1960, p. 16). Cost layers were constructed from the following information on transportation costs. The cost layers for the years 1900 and 1922 are shown in Figures A3.1 and A3.2.

**Head porterage:** With 3.83s per ton mile, we found the lowest cost estimate for the 200km distance from Cape Coast to Kumasi in Gould (1960, p. 16). Curiously, Gould (p. 46) gives an estimate of 5.2s per ton mile for the same trade route from Cape Coast to Kumasi and Kintampo, which lies 165 kilometres north of Kumasi. The Gold Coast Handbook cited costs of 5s per ton mile minimum and 6.25s per ton mile maximum (Maxwell 1923). Tsey's (1986, p. 302) figures for the 3 day journey from Koforidua to Dodowa (30km air line distance) - which included a 2s to 3s wage for "the last day's journey over the awkward Aburi hills" - implies a staggering cost range between 9.5s and 11.5s per ton mile respectively. We assume 5s per ton mile for head porterage on trade routes. Headloading was the only method off the documented transport routes. For moving through space without transport routes we assume a penalty raising headloading costs to 8s. Terrain such as a thick forest cover in the forest region makes travelling more difficult. Moreover, tracks do not run in a straight line - which, however, is assumed by the Path Distance function, because this is the least costly way to move through that cell. Prices of head porterage hardly changed between 1900 and 1922.

**River transport:** Goods were transported on the Ankobra, Tano, and Volta River. A steam launch could travel on the Volta River from Ada 80km inlands to Akuse, where falls and rapids prevented further travel. Transportation costs were low, at 1s per ton mile (Dickson 1969). After Akuse canoes had to be used. Most canoes stopped at Padjae (Gould 1960, p. 46). Prices reported in the colonial reports annual in 1903 imply costs for the journey by canoe from Akuse to Padjae of ca. 3.5s per ton mile (Gold Coast, 1904). The Ankobra River was used by mining firms to transport machinery prior to the railway age. Steam launches could progress from Axim 80km inlands up to Tomento. Canoes could go 80km further up to Nkunsia (Dickson 1969, p. 238). Finally, the Tano river is navigable at a length of 72 km up to Tanoso; canoes operated at a rate of 3.6s per ton mile (Maxwell 1923, p. 233).

**Cask rolling:** Rolling casks was the cheapest "traditional" method to transport cocoa and palm oil springing up at the end of the 19th century. Casks weighed 10-12cwt (ca. 500-600kg) and were rolled by 2-3 laborers (Gould, 1960, p. 25). Roads had to be levelled and of a descending slope; they were few and of short distance linking ports to the hinterland. Cask roads in 1900 include Dodowa-Prampram (30km), Ayimensa-Accra (30 km), Swedru-Winneba (24km), and Mankessim-Saltpond (20km) (Gold Coast 1903; Gould 1960, p. 15). The rate was 1.9s per ton mile.

**Lorry transport 1922:** The quality of roads and the lack of suitable lorries were long an impediment for motor transport. In the first decade of 1900 private transporters expected lorries to last for about three years, with maintenance costs of 25 per cent of the initial value of the lorry per annum (Ntewusu, 2012, p. 97). In 1914 only two lorries were in use (Heap, 1990, p. 22). This changed with the introduction of the light Ford lorries, which were less prone to break down and easy to repair. Lorry imports hugely increased in the late 1910s (Heap 1990). The Gold Coast Handbook put costs in the range of 2.5s and 3.25s per ton mile (Maxwell 1923). For Oda-Saltpond 3.1s per ton mile was charged (Kay and Hymer 1972, p. 249). In the Governor's Annual Address 1923 a rate of 2.75s per ton mile was given (Kay and Hymer 1972, p. 145). We apply the latter figure to roads that fall into the following two categories.

---

1. The practice of cask rolling spread in the 1900s. In 1902 a cask road from Apam to Abodom (30km) was constructed (Gold Coast, 1903, p. 56). In 1905 Dodowa was connected to Ayimensa. After that all export production went to Accra rather than Prampram (Gould 1960, p. 38). In 1907 the Somanya-Akuse road allowed cask rolling (Gould 1960, p. 40). Cask rolling was also prevalent around Kumasi. Cask rolling survived until the early 1920s when it was eventually replaced by lorry transport.
Note: Transportation rates: steam launch (1s), cask rolling (1.9s), canoe (3.5s), head porterage (5s), lorry transport (2.75s), eastern railroad line (0.75s), and western railroad line (0.625s, 0.5s, 0.375 and 0.25s for the 1st, 2nd, 3rd, and 4th 50 miles respectively). For moving through space without transport routes (white area in maps above) we assumed higher costs of headloading (8s).
i) motor roads and ii) roads generally fit for motor transport in 1922, as shown on a 1922 road map published by the government (Rowe and Gold Coast Survey 1922).

**Railroad:** Railroad tariffs varied by commodity, line, and distance. For cocoa, the Eastern Line Tafo-Kumasi charged 0.75s per ton mile throughout. The Western Line Sekondi-Kumasi had differential rates of 0.625s, 0.5s, 0.375, and 0.25s per ton mile for the first, second, third, and forth 50 miles respectively (Maxwell 1923).

**Vertical Cost Factor:** We calculated the surface distance taking into account the somewhat increased distance due to vertical movements. Ascending slopes make transport more costly, e.g. compare costs of headloading of 5s per ton mile with 9.5s to 11.5s per ton mile when crossing the Aburi hills. We added more complexity by adding a penalty for slopes. For slopes greater than 2.5, 5, 7.5, 10 degrees we assumed a penalty of 10%, 25%, 50% and 100% respectively (so for moving through a cell with a 4 degree slope head porterage costs of 8s per ton mile would increase to 8.8s per ton mile). Slopes greater than +/-15 degrees are prohibitively costly and act as barriers that cannot be passed. The parameter choice is on the conservative side. For instance, transporting one ton of cocoa from Koforidua to Dodowa crossing the Aburi hills, cost between £8.55 and £10.35 (Tsey 1986, p. 302). We calculated costs of £7.6 without and £8.85 with the penalty for terrain. We did not impose cost penalties to transport routes (shown in Figure A3.1 and Figure A3.2), as sources often reported total costs from place A to B, which already incorporates terrain difficulties.

**REFERENCES**


---

2The government started a program to improve road infrastructure in 1924: the new "tarmet" (tarred and metalled) roads were suitable for heavy traffic and throughout the year. As a consequence, prices decreased considerably. Gould (1960, pp. 66-68), for example, gives a rate of 1s per ton mile from Sunyani to Kunso and Kumasi, and 0.85s per ton mile from Prasu to Cape Coast. The tarmet program, however, came too late to have influenced cocoa planting in 1922 and production in 1927.

3Railroad rates were lowered in the 1920s. In 1927, for example, rates at the Eastern Line were ca. 0.55s per ton mile (Maxwell 1928, p. 354).
Web Appendix 4: Profit Analysis

1 Transportation Costs and Profitability of Cocoa Cultivation

What are the maximum transportation costs to keep cocoa farming profitable? We calculate this as revenue minus costs excluding transportation costs - expressed in present values (PV) and then determine the transportation costs that would push the NPV to zero.

**Costs:** Cardinall (1932, pp. 86-90) described in detail the costs of cocoa farmers. While his figures refer to the 1930s, his discussion is valid for cocoa farming in 1900 (Austin 2012).

Cardinall (1932) reported that to bring one acre of land to production, it took 115 man days in the first year (clearing, stumping, planting, weeding) and 14 man days in each of the consecutive four years (weeding). Cardinall valued cocoa seedlings at 10s per acre - a price that apparently reflects the very depressed cocoa price during the world economic crisis.\(^1\) We instead assume a price twice as much in the period before 1930.

After the fifth year, cocoa trees bear production. Cardinall (1932, p. 88) reported that labor input then becomes 24 man days per acre per year (weeding, pruning, picking, fermenting). Beckett (1944) described slightly different labour inputs doing more justice to the production cycle of the cocoa tree which does not get into full production until year 10. His estimates are 13.3 man days per year per acre before year 10 increasing to 29.8 man days afterwards. Following Austin (2012) we prefer Beckett's estimate.

We follow Cardinall's figures accounting for depreciation of production tools (cutlasses, hoes, pruning knives, pickers) and drying racks. The former was given at 5s per annum and occur from year one on, the latter were 2.5s per annum and occur once trees produce yields. The costs are not expressed on a per acre basis but refer to the whole farm.

We make two important qualifications to Cardinall's analysis. First, following Austin's argumentation (2012), we set the wage - or opportunity costs in the case of family labour - at 1.25s per day rather than the 1s assumed by Cardinall. Labourers were often provided with land to grow food crops and where this was not the case the wage was supplemented with a subsistence allowance in the order of 25% (Austin 2012). Indeed 1.25s was the going wage for unskilled labour in the Gold Coast (e.g. carriers) - the same category into which agricultural labourers fell. Secondly, Hill (1963) documented that most of the cocoa farmers were actually 'migrant farmers' growing cocoa on land not belonging to their own stool. This affects the costs in at least two respects. First, 'strangers' had to buy land use rights from the chief. Hill (1963, p. 49) reported an average price of 20s per acre in the Nankese area in 1906. This has to be included in the capital outlay to set up the cocoa farm. Second, strangers had to pay a higher rent per bearing cocoa tree. If on own village land, farmers paid 0.25d per tree (ca. 5s per acre) as a contribution to stool funds, whereas strangers paid a rent of 1d per tree (£1 per acre) (Cardinall 1932, p. 87).

We consider the first 5 years as capital formation. Following Cardinall, we take a 5% interest rate as the cost of capital, and assume an average life of a cocoa tree of 30 years. In practice, trees were productive for much longer - about 20 years - though yields per acre decreased from year 20 on (Brecher and Parker 1974). For our purposes, the 30 year life span is adequate - as this was the expectation in the 1900s when planting decisions were made. We assume a linear depreciation of capital (cocoa trees) - starting at £12.94 in year 5 and decreasing at a rate of £0.52 to 0 in year 30.

Cardinall put the average farm size at 4 acres. With the average yield of 9 loads (244.9kg) per acre, 4.08 acres are required to produce one ton.\(^2\) Hence, we multiply costs by a factor of 1.02. We arrive at net present costs of production of £134.7 and £171.7 per ton for 'native' and 'stranger' farmers respectively.

**Revenues:** Prices varied remarkably over 30 years. The average producer price net of export duties was £38.4 per ton (sd: 14) in 1908-1930 and £31.3 per ton (sd: 9.1) in 1920-29 (Viton 1955). We subtract merchant expenses (brokerage, bags, storage, overhead) that Cardinall (1932, p. 89) valued at £2.55 per ton of cocoa. The present value of revenues ranged between £318.6 and £395.9.

\(^1\)According to Cardinall, one acre produced 0.24 tons of cocoa. One tree produced about 6 pods (Elliot 1973). Since seedlings are one alternative use of pods, the price of seedlings must be about 6x0.24 of the price of one ton of cocoa. Based on the producer price in 1931/32, this is about 13s per acre. Cardinall's lower price of 10s per acre can be attributed to the fact that seedlings were not imported but passed on locally and hence transportation costs were lower.

\(^2\)Cardinall's yield estimate is between Becket's estimate of 190kg per acre for Akokoaso and the agricultural department's estimate of 272kg for Ashanti (Austin 2012).
The maximum transportation costs are shown in Table A4.1.

Table A4.1: Maximum transportation costs per ton (£) depending on farmer’s status and price

<table>
<thead>
<tr>
<th></th>
<th>Price of 38.4 per ton</th>
<th>Price of 31.3 per ton</th>
</tr>
</thead>
<tbody>
<tr>
<td>Native Farmer</td>
<td>18.5 (23.9)</td>
<td>13.0 (16.9)</td>
</tr>
<tr>
<td>Stranger Farmer</td>
<td>15.9 (20.6)</td>
<td>10.4 (13.6)</td>
</tr>
</tbody>
</table>

Notes: Based on a 5% discount rate (transport costs for 0% discount rate in brackets).

We can add more complexity taking into account the yield patterns over the full production cycle of 55 years (Brecher and Parker 1974), plus the observed year to year fluctuations in real prices (Frankel 1974). We assume that all production from a given year is sold in that year. The result is shown in Figure A4.1 below.

These are lower bound estimates. There are other costs such as mobility costs when moving from one location to another, building a new house, and losing social networks. We did not add a risk premium, e.g. assuming perfect foresight in prices, though prices are hardly predictable over 30-55 years.

![Figure A4.1 Maximum transportation costs per ton depending on farmer’s status and year of establishment of the cocoa farm.](image)

Notes: Discount rate in parentheses.

## 2 Cocoa Farming vs. Subsistence Farming

How much richer was a cocoa farmer compared to a subsistence farmer? For an estimate of the income of a cocoa farmer we need to take into account transportation costs. Using the map of cocoa production in 1927 and our cost layers (see Web Appendix 3) we calculated average transportation costs of £9.8 (based on transportation methods of 1900 and railroad of 1922).\(^3\) We abstract from issues such as when reinvestments occur and base the analysis on the average income stream in Cardinall’s 30 years horizon. The average annual profit of the native cocoa farmer is

\(^3\)In 1922 rail shipping costs from Tafo (65 miles) and Kumasi (168 miles) to the ports amounted to £2.5 and £4 respectively. Cardinall (1932, p. 89) added transportation costs to his calculation. Based on an average distance of 30 miles and a rate of 0.75s per ton mile he arrived at shipping costs of £1.125 from the farms to Kumasi rail station. For comparison, the average transportation costs to get the 1927 cocoa production to the ports using the 1900 transportation methods would have been £20.7 per ton; using the 1922 transportation methods including motor roads the costs would have been £7.7 per ton instead.
between £13.8 and £6.9. Note that 'stranger farmers' owned larger farms than the 4 acres assumed by Cardinall and so their profits were likely to be much larger (Hill 1963).

We do not know the income of a subsistence farmer, but we can value the time of a subsistence farmer by his opportunity costs (1.25s per day) and multiply this with the days worked per year - a typical assumption being 250 days a year (Allen 2001). This gives a yearly income of £15.3. Note that because we valued labour on the cocoa farm as opportunity costs at 1.25s, substitution of labour between cocoa farming and subsistence farming would have a neutral effect. Thus, the extra income of the cocoa farmer consists of the rent created by capital that cocoa trees represent only. Overall, a cocoa farmer is about 45%-90% richer than a subsistence farmer.

REFERENCES


---

4Frankema and Waijenburg (2011) assumed 312 days per year when estimating urban labour income. This figure appears high given that agricultural labour is seasonal and subsistence farmers do - by definition - not increase labour input to generate surplus production to be sold for markets.
Web Appendix 5: Social Savings

1 Social Savings

Social savings are defined as (Fogel 1964):

\[ \text{Social savings} = (c_a - c_r)R \]

(1)

where \( c_a \) and \( c_r \) represent marginal costs of the railroad and the next-best alternative transport mode respectively; \( R \) is the total volume transported by rail.

We are interested in the social savings of the rail network in 1922. Because cocoa trees start bearing fruits after five years, the impact of the network in 1922 can only be observed in 1927 at the earliest.\(^1\) However, cocoa production in 1927 was transported using the rail network in 1927. Hence, we need to assign cocoa haulage to 1922 stations. We used the ArcGIS Path Distance Allocation function and the cost layer of 1900 transportation with 1922 railroads to determine the least costly path to a port (see Web Appendix 3). Bompata station marks the border where shipment via Kumasi is less costly. We therefore assign all cocoa that was shipped from Bompata and stations to the east in 1927 to Kumasi station. Tafo, the terminus of the Eastern line in 1922, would have received all cocoa hauled from stations west of Bompata in 1927. For the Huni Valley-Kade line, we find that the railroad would only have attracted traffic from stations west of Twifo Prasu - which is just in the 40 km vicinity of the Western line. For the remaining stations, transportation modes of 1900 would have been cheaper. Thus, we ignore the haulage from eight stations, which amounted to 466 tons only. We then derived total freight ton miles \( R \) by taking the product of the quantity of cocoa loaded at each rail station and the distance to the port (Gold Coast 1929). As a result, we find that the railroad network of 1922 would have transported a total of 13,535,228 ton miles of cocoa in 1927.

Under zero profits freight revenues equal costs \( c_a R \). Available evidence suggests that supernormal profits are indeed negligible (Gould 1960).\(^2\) In 1927 lorry transport was a strong competitor to railroads. Freight charges were reduced driving down profits. Nevertheless, relaxing the assumption of perfect competition, allowing for positive profits, would not change conclusions, because in our case social savings are overwhelmingly determined by the very high costs of the alternative transport method.

We then calculate the hypothetical costs, if the same volume of cocoa was moved by the alternative which is head-loading. Our cost estimate of 5s per ton mile is based on what contemporaries reported, which reflects the relatively high wage for free unskilled labour in the Gold Coast colony. The cost difference is the social savings, which amounted to 7.9% of GDP.\(^3\)

\[ \text{TABLE A5.1: Social Savings of Cocoa Production in 1927 using 1922 Rail Network (Cocoa Transport)} \]

<table>
<thead>
<tr>
<th>Social Savings (( £ ))</th>
<th>GDP (%)</th>
<th>Social Rate of Return (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total freight in 1927 on 1922 network (ton miles)</td>
<td>13,535,228</td>
<td></td>
</tr>
<tr>
<td>Total freight revenue (( £ ))</td>
<td>353,322</td>
<td></td>
</tr>
<tr>
<td>Capital outlay in 1922</td>
<td>6,295,990</td>
<td></td>
</tr>
<tr>
<td>Costs of alternative means of transports:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cask rolling (1.9s/ton mile)</td>
<td>1,285,847</td>
<td>932,525</td>
</tr>
<tr>
<td>Head porterage (5s/ton mile)</td>
<td>3,383,807</td>
<td>3,030,485</td>
</tr>
</tbody>
</table>

We also report social savings if cask-rolling is taken as alternative. Especially north of Accra and around Kumasi, cocoa was rolled in casks of about 500 kg (Tudhope 1909: 43). However, cask rolling required suitable roads with a downward slope so that the casks could be moved by muscle power (Dickson 1968, p.38; Hogendorn 1969, p.323). Hence, freight rates did not reflect road maintenance costs and therefore understate the full costs of this means of transport. Moreover, it is doubtful whether geography would have allowed an extensive network of cask roads at low costs. Cask roads were generally short.

\(^1\)Our estimates of the Tafo-Kumasi placebo lines suggest that cocoa growing in expectation of the railroad can be neglected.\n\(^2\)The social savings approach counts economic profits as a cost instead of a rent, thereby underestimating the gain in welfare.\n\(^3\)The GDP figure of £38,290,000 (current £) refers to 1930 and was derived from Omaboe (1960), who revised figures from Cardinal (1931). Chaves (2012), in contrast, arrived at a GDP of £53,108,142 in 1925-26 while Jerven’s (2012) reconstruction attempts suggest a figure close to Omaboe (1960).
TABLE A5.2: Social Savings of Cocoa Production in 1927 using 1927 Rail Network (Cocoa Transport)

<table>
<thead>
<tr>
<th></th>
<th>Social Savings (£)</th>
<th>GDP (%)</th>
<th>Social Rate of Return (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total freight (ton miles)</td>
<td>14,807,649</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total freight revenue (£)</td>
<td>363,910</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Capital outlay</td>
<td>8,432,831</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Costs of alternative means of transports (£)

<table>
<thead>
<tr>
<th>Costs</th>
<th>Social Savings (£)</th>
<th>GDP (%)</th>
<th>Social Rate of Return (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cask rolling (1.9s/ton mile)</td>
<td>1,406,727</td>
<td>1,042,817</td>
<td>2.7</td>
</tr>
<tr>
<td>Head porterage (5s/ton mile)</td>
<td>3,701,912</td>
<td>3,338,002</td>
<td>8.7</td>
</tr>
</tbody>
</table>

For the social rate of return, we relate the social savings to the capital outlay in 1922. We find an astonishingly high social rate of return of 48.1% if headloading is taken as alternative. We also report social savings of 1927 based on the 1927 rail network in the table below. Social savings were 8.7% and the social rate of return was 39.6%.

Figure A5.1: Social Savings and Inelastic Demand for Transport.

Figure A5.2: Social Savings and Elastic Demand for Transport.
2 Social Savings as Consumer Surplus

To see the concordance between social savings and welfare, it is best to think of a simple supply-demand framework of transportation where cocoa farmers represent transport consumers. Under perfect competition transport prices $p$ equal marginal costs $c$, hence horizontal supply curves. Furthermore, a completely inelastic demand for transport is assumed. It can be readily seen from the left figure that under these assumptions social savings are identical to the gain in consumer surplus in the transport market. Jara-Diaz (1986) showed that consumer surplus in transportation exactly equals the net economic benefits created in the products' markets.

However, our context lets us expect that the social savings estimate is much larger than the actual gain in welfare. First, demand was highly elastic (as drawn in the right figure): our results suggest that railroad construction triggered significant cocoa production inland. Second, the railroad was much more efficient than available alternatives in 1900. This cost difference $p$ ranged between 4 and 10 for cask rolling and headloading respectively. The spread between $c_r$ and $c_a$ exacerbates the discrepancy. Overall, assuming $R$ under $c_a$ implies a deadweight loss that the social savings approach counts as a gain (shaded area in the right figure).

REFERENCES


\[^{4}\text{We thank Robert Allen for this point.}\]
Web Appendix 6: Enumeration Area Map of 2000 Ghana Census

The 2000 Facility Census listed 88,969 localities and 12,556 enumeration areas (EA). A locality is defined as a physically distinct settlement (Ghana Statistical Service, 2005). This can be a single house, a hamlet, camp, farm, village, town or city. Population size therefore varies. Unfortunately, the 2000 Facility Census data set lacks information on geographic location.

We retrieved the geographic coordinates of localities by matching place names with several geographic data bases. In a first step, we used i) the Ghana country file from the GeoNet data base and ii) the index of place names to a 1908 1:125,000 map of the Gold Coast Colony (Guggisberg 1908; National Geospatial-Intelligence Agency 2007). The two databases contain i) 17,133 and ii) 5,640 “official” names of populated places as well as alternative spellings representing a total of 12,780 unique places. Matching resulted in 3,131 unique identifications. Alternative spellings not listed in databases i) and ii) were the most important reason for the large number of unmatched places. Patterns of misspellings were too complex to construct a set of specific rules.

In a second step, we calculated Levenshtein distances between the localities in the 2000 Census and the Geographic databases that were yet unmatched. The Levenshtein distance is defined as the minimum number of characters that need to be replaced, inserted or deleted to transform one string (the 2000 Census place name in our case) into another string (the place name from the Geo database). We restricted matches to places that a) started with the same letter and b) where the Levenshtein distance was smaller than 3 and c) where the ratio of the Levenshtein distance to the length of the Census place name was smaller than one third. The latter criterion would allow a Levenshtein distance of 1 and 3 for a place name with 3 and 9 characters in length respectively. The choice of criteria a) to c) helped to reduce type II errors in the matching process (matching places even though the match is wrong). After this, we went through each individual match and checked for agreement in the pronunciation of the place names. In this way, we identified 2,143 localities. Examples of the five biggest places we matched, with the name in the geographic database in parentheses, include Kwashieman (Kwasiman), Pankrono (Pankronu), Ayigya (Ayija), Dunkwa on Ofin (Dunkwa on Offin) and Gumbihini (Gumbehene).

In a third step, we consulted iii) the 1960 Census Enumeration maps and iv) district poverty maps identifying locations by hand (Ghana Census Office 1961; National Development Planning Commission 2012).

We only selected geographic coordinates if they would place the locality within the district boundaries as reported in the 2000 Census. Thus, measurement error will be limited to within the correct district. The population of the
Figure A6.2: Voronoi Map of 2000 Facility Enumeration Areas.
8,421 identified localities represent 79.3% of the population counted in the 2000 Census. We then replaced missing coordinates with the average latitude and longitude of known localities of the same EA. This is unproblematic, especially where EAs are small. For the remaining cases, we linearly interpolated coordinates between neighboring EAs. This can be justified due to the coding system of EAs: EAs run in serpentine order within a district from South to North (see the example in Figure A6.1). Exceptions are larger localities such as cities which have EA codes ending on 000, 400 and 500 - which we identified individually. We interpolated a total of 4,743 EAs representing 14.5% of the population counted in the 2000 Census. For 681 localities amounting to 0.4% of the population we failed to retrieve geographic coordinates.

How accurate is the geographic interpolation? For one district - Suhum Kraboa Koaltar - we consulted a 1:50,000 map and identified at least one locality for each of the 140 EAs adding 50 EAs to the 90 EAs retrieved following steps 1 to 3 above (see Fig. A6.1). We find that 16 of the 50 interpolated EAs would have been attributed to the wrong 0.1x0.1 grid cell. The average error in the location, however, was 3.5 km only. We approximated the map of Census EAs applying a Voronoi tessellation. The Voronoi map is shown in Figure A6.2.

REFERENCES


THE GLOBAL CONVERGENCE OF INCOME DISTRIBUTION

John Luke Gallup
Portland State University
1721 SW Broadway
Cramer Hall, Suite 241
Portland OR 97201
jlgallup@pdx.edu
tel: 503-725-3929
fax: 503-725-3945

September 15, 2012

Abstract

What happens to income distribution during the course of economic development? New higher quality international data show a marked pattern of inequality convergence, where inequality becomes more similar across countries as income levels rise. Inequality has tended to fall in high inequality countries as their economies grow, and inequality has tended to rise in low inequality countries. This is clear in linear regression trends, piecewise trends, and stochastic kernel estimation. The stochastic kernel estimation models the evolution of inequality in the income domain, rather than the more typical time domain, and allows for complex dynamics. The evidence of inequality convergence is confirmed in numerous robustness checks. The pattern of convergence is consistent with both rising inequality in many high-income countries and falling inequality in high inequality developing countries, such as many Latin American countries.

JEL Codes: D31, O15, O47
Keywords: Income inequality, Convergence, Cross-country
Word count: 9,350
1 Introduction

What is the effect of economic growth and development on income inequality within countries? Although there has been a great deal of empirical research on this topic since Kuznets (1955), the question has been difficult to answer due to lack of good quality international data on income inequality.

This paper uses new cross-country time series that have been generated using more consistent definitions and methods than in the past. The data set is the first panel with internally consistent time series for a large number of countries. The new data are graphed in Figure 1, showing income levels and Gini coefficients for 87 countries over the past several decades.

The shape of the graph, like a funnel tipped on its side, suggests that inequality converges and declines as income rises. The range of inequality across countries is much wider at low income levels than at high income levels, and the level inequality tends to be higher at low incomes than at high incomes. A naïve linear fit of the data points in Figure 1 shows the Gini coefficient falling from above 40 to 20.

This paper tests for the two apparent patterns in Figure 1: inequality convergence and inequality decline as income grows. The tests of inequality convergence use three methods: linear trends of inequality with country fixed effects, piecewise linear trends, and stochastic kernel estimation. All three methods show a clear convergence of inequality
levels as income grows. Evidence for a decline in typical inequality as income grows is more ambiguous. The linear trend estimation does not show a decline in inequality, but the stochastic kernel estimation does show a tendency for inequality in low- and midde-income countries to decline.

A huge literature has studied the trend of inequality as income grows, almost all of it evaluating the Kuznets (1955) hypothesis. Kuznets predicted that inequality increases at early stages of economic development until a country reaches middle income levels, and then gradually declines, resulting in an inverted U-shaped path of inequality as income grows. Kuznets made his conjecture on the basis of the scantiest of evidence, but subsequent cross-sectional data appeared to support the regularity: middle income countries typically have the highest income inequality. The problem is that cross-sectional data at one point in time cannot convincingly test a hypothesis about change within countries.

Although the Kuznets hypothesis continues to dominate thinking about the relationship of development and income inequality, even with improvements in data availability, there has never been convincing evidence for the inverted U pattern. Due to space limitations, the Kuznets hypothesis is evaluated using the new data in a companion paper (Gallup, 2011), which also surveys this literature.

The effect of growth on inequality and the effect of inequality on growth are two sides of an endogenous feedback. A substantial literature
has looked at how initial inequality affects subsequent economic growth. Empirical findings have differed, with substantial doubt about the quality of the inequality estimates used. Bénabou (1996) synthesizes the theory behind inequality’s effect on growth as well as the empirical evidence, and Banerjee and Duflo (2003) critique subsequent results. The issue of endogeneity will not concern us here, however, since we are studying the reduced form patterns of their joint evolution, rather than trying to isolate which is causing which.

Data has been a problem for research on the relationship between economic growth and inequality from the beginning. Even a rudimentary cross section of country inequality levels did not become available until more than a decade after Kuznets’s original paper. Income distribution statistics are calculated from household income or expenditure surveys. By now these surveys have been conducted in many countries around the world for a substantial number of years, but until very recently no organization has calculated consistent inequality series from the survey data for a large number of countries.

Deininger and Squire (1996) assembled the first large-scale dataset with enough observations to study the typical path of inequality within countries. They took the estimates of Gini coefficients from hundreds of separate studies of inequality in individual countries to construct a large number of country time series. Inequality measures depend sensitively on the definition of income or expenditure, unit of observation, survey
coverage, etc. Atkinson and Brandolini (2003) were skeptical of the accuracy of time series constructed by linking inequality estimates from separate studies of inequality using potentially inconsistent definitions over time within each country. They showed that for a large number of Western European countries, Deininger and Squire's time series departed significantly from an independent series they constructed using consistent definitions. Western European data should presumably be among the most accurate. Atkinson and Brandolini's results cast a shadow over the research using the Deininger and Squire dataset.¹

All the country time series used in this study are calculated from repeated rounds of the same national household survey using the same definitions and statistical methods throughout, which was not feasible at the time Deininger and Squire constructed their dataset.

The main innovations of the paper are to compile a more reliable panel of inequality time series than available in the past, and to present evidence of a pattern of inequality convergence. The paper is empirical, with only brief speculation about the possible causes of inequality convergence in the conclusion.

The studies most similar to this paper are two brief analyses of the

¹ A separate line of inquiry has compiled much more accurate data on the share of income earned by the richest individuals, or “top incomes”. The income shares are calculated from income tax records, and provide very long time series for countries with a long history of income taxation. So far these data are available for a limited number of mostly high-income countries which makes them unsuited for evaluating a general convergence of inequality. Atkinson and others (2011) survey the progress of this research so far.
convergence of cross-country inequality by Bénabou (1996) and Ravallion (2003), using inequality data from Deininger and Squire (1996).\(^2\) Both studies evaluate inequality convergence over time rather than convergence as income levels rise, which is misspecified if inequality changes due to economic growth. Bénabou uses a test of “sigma convergence”, finding ambiguous results, and Ravallion uses a test of “beta convergence”, finding evidence of convergence. Quah (1993a, 1993b) has criticized both sigma convergence and beta convergence as tests of convergence in distribution, and as already noted, the reliability of time trends in Deininger and Squire’s data has been called into question (Atkinson and Brandolini, 2003). Quah (1997, 2007), studying cross-country convergence of income levels, proposes stochastic kernel estimation as a more direct test of convergence in distribution, which is adopted by this paper.

Some studies have tested for the convergence of inequality between states within a country. Kar and others (2010) apply stochastic kernel estimation to study convergence of inequality between Indian states over time since economic liberalization in the early 1990s. They find a divergence of state inequality over this period. Panizza (2001), Ezcurra and Pascual (2009), and Lin and Huang (2009) study convergence of inequality of U.S. states over time, and find a strong tendency towards convergence. Ezcurra and Pascual use stochastic kernel estimation, and

\(^2\) Ravallion (2003) also uses a separate inequality dataset, which is not as subject to the time series problems of the Deininger and Squire dataset, but this dataset includes only 21 countries.
Panizza and Lin and Huang use variants of the beta convergence test.

Bénabou (1996, p. 58) follows his analysis of the cross-country convergence of inequality with these thoughts: “My main purpose ... was to put forward the issue of convergence in distribution as an important and essentially unexplored topic for empirical research. Hopefully, future studies with more sophisticated econometrics and better data ... will help resolve the issue.”

Bénabou gave three reasons in favor of studying whether countries are converging to the same level of inequality. “First, ascertaining the facts is itself of interest. Second, it can shed light on the relevance of models with multiple steady states and history dependence of the income distribution. Third, ... [o]nce augmented with idiosyncratic shocks, most versions of the neoclassical growth model imply convergence in distribution (see for instance Banerjee and Newman (1991), Ayagari (1994), Bertola (1995), or the case without credit constraints in Piketty (1997)).”

The rest of the paper is organized as follows. The next section discusses the new inequality data. Section 3 considers the path of inequality using linear regression with country fixed effects and piecewise linear regression. Section 4 measures change in the whole distribution of inequalities using stochastic kernel estimation, providing a prediction of where inequality is headed. Section 5 discusses some country and regional trends. Section 6 concludes.
2 Data

The lack of comparable inequality data across countries has always plagued research on income inequality. Unlike most other basic national statistics, there is no single international organization that collects measures of income distribution worldwide.

Until the mid-1990s, only cross-sectional country data were available, for a limited number of countries, and they were extracted from many different studies of uneven quality. Deininger and Squire (1996) dramatically increased the size of the collection and classified the studies by quality, obtaining enough studies to construct time series for many countries. However, as noted by Atkinson and Brandolini (2003), combining estimates from separate studies using different methods and survey coverage into a time series is problematic because the level of inequality estimated is quite sensitive to the details of how the data are collected and how the inequality statistics are calculated. The apparent time series trends may simply reflect the idiosyncrasies of the separate studies which make up the series.

Only very recently have consistently constructed time series of income distribution become available for a large number of countries. Four regional organizations have created series of inequality statistics spanning more than a decade: Eurostat (2011) for European Union
members, the TransMONEE database created by UNICEF for Eastern Europe and former Soviet Union countries (TransMONEE, 2011), SEDLAC (2011) for Latin America and the Caribbean, and the Luxembourg Income Study database (LIS, 2011) for other high income countries. Together these four organizations provide statistics for all of Europe (East and West), Central Asia, Latin America, and most other high-income countries.

Major parts of the world still do not have standardized collection of income distribution statistics: East and South Asia, the Middle East, and Africa. Of these, only the Asian regions have a large number of countries with the raw material for the statistics: household income and/or expenditure surveys spanning a substantial number of years.

In addition to statistics from the four regional organizations above (Eurostat, TransMONEE, SEDLAC, and LIS), this study uses statistics from other countries where there are time series of consistent quality from national statistical agencies and academic studies. Many of these statistics were compiled in UNU-WIDER World Income Inequality Database (WIID 2011). WIID is the culmination of a series of previous efforts to compile secondary data on income distribution, building on the work of Deininger and Squire (1996, 1998) and earlier collections described therein. Unlike Deininger and Squire, however, this study only includes data from countries where there is a unified time series using the same survey and the same methodology throughout.
The country times series used in this study hew to the following criteria:

1) They are calculated from surveys of household income (covering all income sources) or household consumption expenditure drawn from a national sample of all households.

2) The time series are calculated from surveys with the same survey design each year.³

3) The time series of inequality statistics are calculated using the same method and definitions throughout.

The lack of consistent definitions Deininger and Squire's inequality time series in point 3) is what caused doubt about their accuracy. Most of the data from Eurostat, TransMONEE, and SEDLAC are derived from national surveys established in the 1990s, so they were not available when Deininger and Squire compiled their data. Even most Western European countries did not have annual household income surveys before the establishment of a European-wide survey in 1995.

All the inequality statistics used in this study are Gini coefficients of household income or household expenditure. Distribution in low-income countries is usually measured by inequality of household consumption rather than income because many workers are self-employed, and the self-

³ Minor changes of survey questions and survey design still occur over time in many of the standard national surveys, but statisticians usually address these inconsistencies when they calculate a time series of inequality estimates.
employed often do not distinguish their net income from their revenues. Distribution in high-income countries, as well as Latin American countries, is measured by inequality of household income because it mostly derives from easily recalled wages, and few workers are self-employed.

An online data appendix documents the source of each country's inequality data, years of coverage, and method of measurement. Differences in the income concept used (household consumption, household disposable income, or household gross income) do usually affect the calculated level of inequality (Deininger and Squire, 1996, Coulter, Cowell, and Jenkins, 1992, and Jenkins and Cowell, 1994). The weights applied to different household members are less likely to have a major impact on the level of inequality. Since the regression analysis in this study controls for different country levels of inequality, the priority is accurate measurement of inequality change over time rather than precisely comparable inequality levels. Neither differences in income concept nor differences in household weights are likely to have much impact on the measurement of change in inequality.

The data include time series from 87 countries, split regionally so that about a quarter of the countries come each from the OECD90\(^4\), Latin

\(^4\) OECD90 indicates members of the Organization of Economic Cooperation and Development as of 1990, which includes the highest income countries of the world except for oil exporting countries. More recent OECD members Mexico, South Korea, Chile, Israel, Czech Republic, Slovakia, Estonia, Hungary, Poland, and Slovenia are excluded from the OECD90 group, since most of these countries still have income levels substantially lower than OECD90 members. The excluded countries are included in the other regional groups.
America, Eastern Europe and the Former Soviet Union, and Asia and Africa combined, as shown in Table 1. Asia and especially Africa are underrepresented and tend to have shorter time series. The country time series have an average span of 15.3 years, and an average of 9.9 observations per country. The number of observations per country varies from 2 to 32. The average frequency of the observations is every 1.6 years.

Table 1. Regional coverage of data

<table>
<thead>
<tr>
<th>Region</th>
<th>No. of countries</th>
<th>Percent of countries</th>
<th>Percent of observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>OECD90</td>
<td>24</td>
<td>28</td>
<td>35</td>
</tr>
<tr>
<td>Latin America</td>
<td>19</td>
<td>22</td>
<td>25</td>
</tr>
<tr>
<td>Eastern Europe and the</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Former Soviet Union</td>
<td>21</td>
<td>24</td>
<td>23</td>
</tr>
<tr>
<td>Africa</td>
<td>5</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>Asia</td>
<td>18</td>
<td>21</td>
<td>16</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>87</strong></td>
<td><strong>100</strong></td>
<td><strong>100</strong></td>
</tr>
</tbody>
</table>

Most of the data (78%) come from 1995 to 2009. Almost all the data are from 1985 or later (91%), but a few countries have good quality data earlier than this, in the case of India going back annually to 1951.

The data for Eastern Europe and the former Soviet Union before 1994 are excluded to avoid picking up the sudden inequality changes after the collapse of central planning. Including these data would strengthen the apparent convergence of income distributions across countries (since
these countries had very low inequality during the Communist era), but were not included because this historical break was not the result of the typical process of economic development observed in other countries.

Gini coefficients are paired with income levels in each country and year. Income levels are measured by gross domestic product (GDP) per capita from the Penn World Tables version 7.0 (Heston and others, 2011). The GDP per capita figures are adjusted for purchasing power parity and reported in 2005 constant international dollars (which are calibrated to represent the purchasing power of one dollar in the United States in 2005).

3 Linear and Piece-wise Linear Trends

The simplest way to evaluate the path of inequality as income grows is with a linear trend. Countries may have higher or lower base levels of inequality for historical reasons or due to differences in measurement, such as calculating inequality using household consumption instead of household income. Since we are interested in the typical pattern of change within countries, we control for different inequality levels in each country with fixed effects estimation.

The ordinary least squares linear fit of the data in Figure 1 suggests a dramatic decline of inequality levels as incomes rise. The pattern of inequality seen in the data could arise even in a world where inequality never changes in any country, as shown in Figure 2. Fixed effects
regression, on the other hand, would correctly estimate that inequality is not related to income level.

Figure 3 and Table 2 show linear and quadratic fixed effects trend lines for all the countries. The trend lines in Figure 3 are superimposed on the individual country data series. The slope of the overall trend is slightly positive, but not significantly different from zero. It is very different from the steep decline in inequality suggested by the ordinary least squares fit in Figure 1.

<table>
<thead>
<tr>
<th>Table 2: Fixed effects trend lines</th>
</tr>
</thead>
<tbody>
<tr>
<td>GDP per capita (‘000 $)</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>GDP per capita²</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Constant</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>$R^2$</td>
</tr>
<tr>
<td>$N$</td>
</tr>
</tbody>
</table>

* $p<0.05$; ** $p<0.01$

Robust $t$ statistics in parentheses

The trend in inequality is not significantly curvilinear. The quadratic trend is virtually indistinguishable from the linear trend and slightly concave, contrary to the inverted-U shape hypothesized by Kuznets (1955).

A simple way to check for convergence of inequality across
countries is to see if inequality levels are more likely to be falling if the country starts out above the trend line of inequality and more likely to be rising if the country starts out below the trend. Figure 3 and Table 2 show the fixed effects trends for higher inequality countries (“above trend”) and for lower inequality countries (“below trend”).

The high inequality “above trend” countries show a large and statistically significant decline in inequality, indicating that high inequality countries have experienced more reduction in inequality in recent decades than lower inequality countries.

The low inequality “below trend” countries have a statistically significant increase in inequality as income grows. If countries start out above the trend their inequality declines relative to trend, and if they start out below trend, their inequality increases relative to trend.

The striking pattern in Figure 1 of apparent convergence of inequality is displayed in terms of the level of GDP per capita. Inequality is more often compared to the natural logarithm of income than to income itself, but in this case convergence is not so apparent to the eye. Using the logarithm of GDP per capita has two advantages. First, if a country is growing at a constant rate, the income will advance by the same

5 The level of the intercept of the predicted fixed effects trend line is set to pass through the median of the Gini coefficients rather than the mean in Figure 3 (and Figure 4). This causes the groups of countries “above trend” or “below trend” to have approximately equal numbers of observations. The exact level at which the overall trend is set has little effect on the estimated “above trend” and “below trend” coefficients.
increment each period on a logarithmic scale. Second, the distribution of
countries by income level is highly skewed so that there are many more
low-income countries than high-income countries. The country data are
more uniformly spaced along a logarithmic scale rather than being
bunched towards the origin.

Figure 4 and Table 3 show the trend of inequality using a
logarithmic scale. The overall trend on the logarithmic scale is slightly
positive and not statistically significant, like the non-logarithmic case.
Similarly, the quadratic term in income is not statistically significant and
the quadratic trend line is almost indistinguishable from the linear trend,
although there is the slightest concavity.

<table>
<thead>
<tr>
<th></th>
<th>Linear trend</th>
<th>Quadratic trend</th>
<th>Above trend</th>
<th>Below trend</th>
</tr>
</thead>
<tbody>
<tr>
<td>ln(GDP per capita)</td>
<td>0.233</td>
<td>-2.749</td>
<td>-2.770</td>
<td>5.038</td>
</tr>
<tr>
<td></td>
<td>(0.25)</td>
<td>(0.36)</td>
<td>(2.95)**</td>
<td>(3.49)**</td>
</tr>
<tr>
<td>ln(GDP per capita)^2</td>
<td>0.170</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.41)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>34.247</td>
<td>47.139</td>
<td>67.194</td>
<td>-20.416</td>
</tr>
<tr>
<td></td>
<td>(3.94)**</td>
<td>(1.33)</td>
<td>(8.26)**</td>
<td>(1.44)</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.00</td>
<td>0.00</td>
<td>0.06</td>
<td>0.18</td>
</tr>
<tr>
<td>N</td>
<td>865</td>
<td>865</td>
<td>452</td>
<td>413</td>
</tr>
</tbody>
</table>

* p<0.05; ** p<0.01
Robust t statistics in parentheses

There are again clear indications of convergence in inequality in this
case. Countries starting out with inequality above the linear trend show a significant decline in inequality. Below trend countries show an even steeper, statistically significant, increase in inequality.

The fixed effects estimates imply that for a typical country starting above the trend level of inequality, when income doubles the Gini coefficient declines by about 1.9 points. A typical country starting below the inequality trend would tend to see an increase the Gini coefficient of 3.5 points when income per capita doubled.

Do the trends in inequality differ for low-income, middle-income, and high-income countries? Is there consistent evidence of convergence throughout the range of income levels? Figure 5 and Table 4 present three different fixed effects trends according to income level, accompanied by the trends of “above trend” and “below trend” countries in each income category. The trend lines by income level in Figure 5 are superimposed on individual country trend lines (rather than the data).

All three of the above trend coefficients are negative and all three of the below trend coefficients are positive, though not all are statistically significant. The above trend coefficient in each income level is, however, statistically significantly different from the below trend coefficient in all

---

6 A doubling of income is an increase in the log of income of 0.693. Applying the “above trend” coefficient for log GDP per capita, $0.693 \times -2.770 = -1.920$.

7 Each fixed effects regression includes an intercept which is not reported. The income categories are each 1.6 log GDP per capita wide, with low income ranging from $590 to $2700, middle income from $2700 to $13,400, and high income from $13,400 and above.
income categories. The differences between the above and below trend coefficients are robust to variation in the income level break points (not shown).

Table 4: Piecewise linear trend lines
log GDP per capita coefficients

<table>
<thead>
<tr>
<th></th>
<th>Low income</th>
<th>Middle income</th>
<th>High income</th>
</tr>
</thead>
<tbody>
<tr>
<td>Above trend</td>
<td>-3.700</td>
<td>-6.627</td>
<td>-1.746</td>
</tr>
<tr>
<td></td>
<td>(1.91)</td>
<td>(3.41)**</td>
<td>(0.98)</td>
</tr>
<tr>
<td>Trend</td>
<td>1.297</td>
<td>-0.979</td>
<td>0.392</td>
</tr>
<tr>
<td></td>
<td>(0.52)</td>
<td>(0.73)</td>
<td>(0.30)</td>
</tr>
<tr>
<td>Below trend</td>
<td>5.770</td>
<td>1.655</td>
<td>3.764</td>
</tr>
<tr>
<td></td>
<td>(1.51)</td>
<td>(1.02)</td>
<td>(3.44)**</td>
</tr>
<tr>
<td>Above - Below</td>
<td>-9.470</td>
<td>-8.282</td>
<td>-5.510</td>
</tr>
<tr>
<td></td>
<td>(2.30)*</td>
<td>(3.31)**</td>
<td>(2.67)**</td>
</tr>
</tbody>
</table>

N  122  357  383

* p<0.05; ** p<0.01
Robust t statistics in parentheses

The robustness of the linear inequality trends is checked in two ways, first using an independent data sample and second with controls for possible measurement error.

The Deininger-Squire (1996) dataset provides independent, if error-prone, estimates of inequality. The statistics cover an earlier period, with no data in common with the sample otherwise used in this paper. A small number of country-year pairs overlap in the two data sets, but each of these data points come from different sources. Few countries overlap in
time. The average first year in the Deininger-Squire data is 1972 and the average final year is 1988, versus an average span from 1991 to 2006 in the new data.

The Deininger-Squire data exhibit the same pattern of convergence, but as one might expect from data subject to substantial measurement error, the levels of significance are a lot lower. Table 5 shows the “above trend” inequality falling in high inequality countries and the “below trend” inequality rising in low inequality countries, but neither estimate is significant at the 5% level. The above trend coefficient is statistically different from the below trend coefficient at the 14% level.

<table>
<thead>
<tr>
<th></th>
<th>Above trend</th>
<th>Below trend</th>
</tr>
</thead>
<tbody>
<tr>
<td>ln(GDP per capita)</td>
<td>-1.902</td>
<td>0.677</td>
</tr>
<tr>
<td></td>
<td>(1.30)</td>
<td>(0.73)</td>
</tr>
<tr>
<td>Constant</td>
<td>58.304</td>
<td>24.119</td>
</tr>
<tr>
<td></td>
<td>(4.82)**</td>
<td>(3.07)**</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.02</td>
<td>0.01</td>
</tr>
<tr>
<td>$N$</td>
<td>317</td>
<td>289</td>
</tr>
</tbody>
</table>

* $p<0.05$; ** $p<0.01$

The new sample, although more consistently measured than the Deininger-Squire data, is surely still subject to measurement error. If there is a large bias in the first inequality estimate in each country, and
the bias diminishes with each subsequent observation, this could generate a spurious pattern of inequality convergence. On the other hand, measurement error that is constant over time, or uncorrelated with time, even if it were high, would not give the appearance of convergence.

The data in this study are all generated from ongoing national household sample surveys conducted by government statistical offices. Almost all of the statistical offices have conducted many other surveys (such as population censuses and price, labor force, and enterprise surveys) for years before undertaking their first household income or expenditure survey. If there is a substantial learning curve for conducting a household economic survey, it is likely to be an issue the first or second time the national survey is fielded. After that the statistical office has established its methods and practice. Some level of error inevitably remains, but there is no compelling reason why it would continue to decline over time. As a point of comparison, estimating income inequality from household economic surveys is much simpler than producing national accounts and GDP estimates, which depend on data from a myriad of different surveys and reporting mechanisms.

To test for measurement error bias, we repeat the estimation of above trend and below trend inequality, omitting the first or the first two inequality observations in each country.\(^8\) Table 6 contains the coefficient

---

\(^8\) Regressions using non-logarithmic GDP per capita (not shown) exhibit the
estimates for income when the Gini coefficient is regressed on the log of income and a country-specific constant. When one or two initial observations are dropped, the results show no qualitatively different patterns for convergence. Above trend countries still show a statistically significant decline in inequality and below trend countries still show a statistically significant increase in inequality. The estimated strength of the convergence is actually slightly greater when the first two observations in each country are deleted. The results reject measurement error as a spurious source of convergence.

Table 6: Measurement error regressions
Fixed effects log GDP per capita coefficients

<table>
<thead>
<tr>
<th></th>
<th>All data</th>
<th>W/o first observation</th>
<th>W/o first two observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Above trend</td>
<td>-2.770</td>
<td>-2.365</td>
<td>-3.070</td>
</tr>
<tr>
<td></td>
<td>(2.95)**</td>
<td>(2.03)*</td>
<td>(2.42)*</td>
</tr>
<tr>
<td>Trend</td>
<td>0.233</td>
<td>-0.482</td>
<td>-0.532</td>
</tr>
<tr>
<td></td>
<td>(0.25)</td>
<td>(0.54)</td>
<td>(0.55)</td>
</tr>
<tr>
<td>Below trend</td>
<td>5.038</td>
<td>3.140</td>
<td>5.151</td>
</tr>
<tr>
<td></td>
<td>(3.49)**</td>
<td>(2.30)*</td>
<td>(3.37)**</td>
</tr>
</tbody>
</table>

* p<0.05; ** p<0.01
Robust t statistics in parentheses

The individual country trend lines exhibit substantial diversity in inequality change. Although the cross-country trend lines indicate overall tendencies, they ignore the churning of inequality levels in different same patterns.
directions at the same income level. The stochastic kernel estimation in Section 4 will allow us to evaluate the consequences of this multidirectional churning of the cross-country distribution of inequalities as income changes.

4 Stochastic Kernel Estimation

The previous section used linear trends in inequality (controlling for country fixed effects) to make a simple evaluation of the convergence of inequality as income grows. A fuller assessment would consider the evolution of the whole distribution of inequalities as income rises, rather than just the linear trend. A linear trend, for instance, could mask a dynamic where certain countries tend towards one level of inequality and other countries tend towards a different level of inequality, even among high inequality countries. The data show substantial diversity across countries in the path of inequality whether they start out with high inequality or low inequality. There are also changes of the direction of inequality within countries, where falling inequality switches to rising inequality and vice versa.

Stochastic kernel estimation is flexible enough to model the net outcome of complex dynamics of inequality change. It can incorporate a dynamic path where all countries tend towards a single level of inequality or where they tend towards two or more different levels of inequality. It can also incorporate a dynamic path where countries switch places.
between different levels of inequality, while the overall distribution of cross-national inequalities remains the same.

Stochastic process models are typically applied to processes that evolve over time. However, most hypotheses about income distribution concern the path of inequality as income grows, not as time passes. For this reason, we model the cross-country distribution of inequalities in the income domain rather than in the time domain.⁹

The stochastic kernel model represents the evolution of continuous distributions from period to period. It is the continuous analogue of a Markov chain, which represents the evolution of discrete distributions. Quah (1997, 2007) explains stochastic kernel models and applies them to the distribution of income levels across countries over time. Since stochastic kernel models use continuous distribution functions, Quah defines them using measure theory instead of the more accessible matrix algebra of Markov chains.

Continuous income distributions are attractive conceptually, but in practice, stochastic kernel models are estimated by discrete approximation. Digital computers must approximate continuous transition surfaces with

---

⁹ Other estimations of convergence of inequality (Ezcurra and Pascual, 1999, for U.S. states and Kar and others, 2000, for Indian states) specify a stochastic kernel in the time domain rather than the income domain. A stationary distribution in the time domain requires assuming that the transition probabilities are constant over time regardless of what is happening to income. Growth theory often suggests a relationship between inequality and growth (e.g. Banerjee and Newman, 1991) but not predictable changes in inequality over time whether or not there is income growth.
discrete grids, so a stochastic kernel model is actually estimated as a fine-grained Markov chain. Since the model is ultimately estimated using a discretized distribution, we will present stochastic kernel estimation using the simpler Markov chain notation. There is no loss of generality since a continuous distribution can be arbitrarily well approximated by a high dimension discrete distribution. Refer to Quah (1997, 2007) for the continuous formulation.

The difference in practice between the stochastic kernel and Markov chain estimation is the method of calculating the transition matrix. Markov transition matrices are typically estimated from crude frequency counts of transitions. Stochastic kernel estimation, in contrast, typically uses bivariate kernel density.\(^\text{10}\)

The stochastic kernel estimation of the transition matrix can be finer-grained with more rows and columns because the kernel density estimation uses information from neighboring cell frequencies to smooth the density estimates. Whereas a typical Markov transition matrix of frequencies from a sample of several hundred observations might be a 5x5 matrix to avoid small sample sizes in each transition cell, a bivariate

---

10 Although stochastic kernel estimation and kernel density estimation both include the term “kernel”, meaning distribution function, they are referring to different uses of a distribution function. Stochastic kernel estimation refers to the estimation of the transition from one period's distribution to the next period's distribution of the variable of interest (here, inequality). Kernel density estimation refers to the weighting scheme for averaging neighboring frequency observations. The weights decline moving away from the cell of interest according to the value of a chosen kernel, or distribution function (e.g. Gaussian, Epanechnikov, rectangular, etc.)
kernel density estimator would commonly produce a 50x50 matrix smoothly approximating the surface.

So in practice a stochastic kernel estimation is a Markov chain estimated with a smoothed transition matrix.

First divide country inequality into $N$ possible levels, with values $g_1, \ldots, g_N$. The set $G = \{g_1, \ldots, g_N\}$ is the state space, and countries move from one inequality level, $g_i$, to another, $g_j$, at each step of income.

A fundamental assumption of the Markov model is the Markov property. Let $X_s \in G$ be the inequality level of a country at income level $s$. The Markov property is that assumption that $E(X_{s+1} | X_0, \ldots, X_s) = E(X_{s+1} | X_s)$. That is, the inequality level at the next higher income level depends only on the inequality level at the current income level, and not on the earlier history of inequality levels at lower income levels. In addition, we assume homogeneity of transitions across income levels: $E(X_{s+1} | X_s) = E(X_s | X_{s-1})$ for all $s$. Then we can define the transition probability as $p_{ij} = E(X_{s+1} = g_j | X_s = g_i)$. The $N$ by $N$ matrix of all the transition probabilities is $P = [p_{ij}]$.

Let $u_s$ be an $N$-dimensional probability row vector which represents the state of the Markov chain at income level $s$. The $i^{th}$ component of $u_s$ represents the probability that the chain is at inequality level $g_i$. Then $u_{s+1} = u_s P$. If we assume that for every inequality level $g_i$
except for $g_i$, there is a positive probability of inequality falling to $g_{i-1}$ at the next income level, and at every inequality level $g_i$ except for $g_N$ there is a positive probability of inequality rising to $g_{i+1}$ at the next income level, these are sufficient conditions for the Markov chain to be *ergodic*, which means there is a possibility of going from every inequality state to every other inequality state, although not necessarily in one step. By Doeblin's Theorem (Stroock, 2000, p. 28), ergodic Markov chains tend towards a unique stationary probability vector as income levels increase without bound. The stationary probability vector $\mathbf{w}$ is defined by

$$\mathbf{w} = \lim_{s \to \infty} \mathbf{u}_0 \mathbf{P}^s.$$  

$\mathbf{w}$ shows us where the distribution of inequalities will end up if the current distributional dynamic continues indefinitely. The stationary distribution $\mathbf{w}$ is equal to the first left eigenvector of the transition matrix $\mathbf{P}$, and is independent of the initial distribution $\mathbf{u}_0$ (Theorem 8.6, p. 106, Billingsley, 1979).

This Markov model allows for a broad range of inequality dynamics such as convergence towards single or multi-peaked inequality levels, divergence, and churning of countries across inequality levels whose overall dispersion does not change.

Estimating the transition matrix requires some conditioning of the data. We need to observe the change in inequality across regular income level intervals. Country inequality levels are measured at regular time intervals (every year in countries with an annual income survey), not at
regular income intervals. To construct a progression of inequality over equally spaced log income intervals, income is quantized into 200 equal log income levels. If more than one sequential inequality observation falls within a given income interval (when income grows too little to progress to the next income level), the inequality observations are averaged.

The categorization of the data into regular income levels produces an inequality “income series” (as opposed to a time series) with a lot of gaps, especially in countries with rapid economic growth, because income may jump several levels between inequality observations. These gaps are bridged by linear interpolation of between income levels in a given country to maintain a connected series (just as a country’s observations are typically connected by straight lines in graphs, as in Figure 1). To prevent rapidly growing countries with more interpolated observations having disproportionate influence, the observations are weighted by the number of actual, non-interpolated, data points when estimating the transition matrix.

The regridding of inequality takes the original 861 observations and interpolates them up to 2,035 inequality transition observations.

The transition matrix P is estimated from a bivariate kernel density, using a Gaussian kernel with a bandwidth of 2.5. The estimation generates a 50 by 50 transition matrix which smooths the raw transition matrix frequencies. The transitions are from the level of the Gini coefficient at the current income level to the level of the Gini coefficient at
the next income level.

The density contours of transition matrix estimated from the sample are graphed in Figure 6. There is a unique peak at point A (a Gini coefficient of 31.5), indicating convergence of inequality levels toward a single peak of attraction.

To the left of point A, the probability mass of the transition matrix is distended to the west, and to the right of point A, the probability mass is distended to the east, drawing inequality levels to the point of attraction at A. For inequality at income level \( s \) on the horizontal axis, consider starting at a Gini coefficient of 22. The density of the transition probabilities is heavily weighted towards higher inequality at income level \( s+1 \). The density mass for the current Gini of 22 is predominantly above the 45° line, meaning that inequality is more likely to rise than fall next income level. In contrast, the density of inequalities next income level starting at a Gini of 38 is mostly below the 45° line, showing that inequality is most likely to fall next period. These tendencies in the transition matrix draw inequalities towards point A as income grows.

The transition matrix \( \mathbf{P} \) shown in Figure 6 has a stationary inequality distribution \( \mathbf{w} \). Figure 7 displays the stationary kernel along with the distribution of inequalities for low-income countries with income per capita below $10,000, and the distribution for high-income countries with income per capita above $20,000. The stationary kernel is more compact than the low-income distribution, but less compact than the high-
income distribution.

The dotted lines indicate the mean inequality for each distribution. The mean inequality of the stationary kernel, at 35.6, is significantly smaller than the mean inequality of low income countries, at 42.3, and significantly larger than the inequality of high income countries, at 29.2.\textsuperscript{11}

The shape and location of the stationary kernel are robust to the choice of kernel and bandwidth. The stationary kernel has a very similar shape using either an Epanechnikov or a rectangular kernel (the latter being about as different from a Gaussian as one can get), although since these kernels limit the range of smoothing, a smoothness similar to the Gaussian kernel requires a higher bandwidth. All three kernels result in a somewhat more compact stationary distribution with a lower mean inequality at lower bandwidths and a more spread out stationary distribution with a higher mean inequality at high bandwidths. A Gaussian kernel with a bandwidth of 0.2 produces a stationary distribution with a mean inequality of 34.3 and a bandwidth of 5 produces a mean inequality of 36.1, compared to a mean of 35.6 with a bandwidth of 2.5. Bandwidths below the chosen level of 2.5 cause local jaggedness in the stationary distribution.

It is possible that inequality change during periods of income decline is different from changes when income grows. To evaluate its

\textsuperscript{11}We can use an ordinary \textit{t} test, since the stationary distribution \textbf{w} is independent of the initial distribution \textbf{u}_0. The means are different at the 1\% level.
impact, the few periods of income decline are removed. At each point where a negative growth observation is deleted, the country series is split into two separate growth spans. Removing negative growth episodes drops 66 observations along with 14 more for new spans that end up being only one observation long, reducing the sample by 9% (80/861). The resulting estimated transition probabilities and stationary kernel are virtually unchanged (not shown), suggesting that inequality changes more or less symmetrically during income decline and income growth.

As with the linear trends in Section 3, the stochastic kernel estimation shows convergence of inequality as income rises. The stochastic kernel estimation finds a modest tendency for inequality to decline from the levels of low-income countries, unlike the results of the linear trend estimation.

One of the basic assumptions of the stochastic kernel estimation is that the probability of change in inequality, given the level of inequality, is the same at all income levels, and hence the stationary distribution to which low-income and high-income countries are approaching is that same. To check this assumption, I split the sample into low- and high-income groups, and vary the cutoff point. This exercise shows a strong bifurcation at a cutoff point of about $15,000 GDP per capita. Low-income countries are strongly convergence to very low inequality levels, and high-income countries converge to somewhat higher inequality levels. The stationary distributions for each of these groups of countries are
graphed in Figure 8. The stationary distribution of inequality for countries with GDP per capita below $15,000 has a mean of 27.1, significantly below even the mean of current high-income countries, and the stationary inequality distribution is tightly bunched near the mean. The means that the transition dynamic is leading all low-income countries towards very low inequality, with little difference between the countries.

The stationary distribution for high-income countries has a higher mean of 32.1 and is more spread out than the stationary distribution for low-income countries.

The shape of the stationary distribution for low-income countries is virtually unchanged if the cut-off for low income is set at $10,000 or at $5,000, but the tendency to converge to low inequality levels diminishes as the cut-off rises above $15,000. For high income countries, the stationary distribution loses its tendency to convergence towards higher inequality at cut-offs below $12,000 or above $18,000.

The results suggest two convergence dynamics, or at least convergence intensities, a powerful one for low-income countries towards very low inequality (even though many low-income countries currently have quite high inequality) and a weaker convergence for high-income countries towards slightly higher inequality than they currently have. The high-income countries are nonetheless converging towards a lower inequality than the current average across all countries.
5 Some country and regional inequality trends

It is a common perception, if unsupported by data, that inequality is worsening around the world. Much of that perception comes from the recent experience of the United and China, the biggest countries in the world in terms of economic size and population, respectively. Both countries have seen a dramatically worsening of inequality in recent decades, but they are uncharacteristic of most of the rest of the countries in the world. It is useful to compare the inequality change in these countries with others.

The U.S. has had rapidly worsening inequality since the late 1970s (Piketty and Saez, 2003, Gottschalk and Danziger, 2003). Many other high-income countries have also experienced modestly increasing inequality over this period, especially in the past decade, but the speed of increase is highest in the U.S.\(^\text{12}\), and its level of inequality is now is much higher than other high-income countries. Figure 10 shows the inequality paths of high-income countries including the U.S. The trend line is the fixed effects trend for countries with GDP per capita above $25,000. The U.S. is now the only high-income OECD country with inequality above the world average.

China, two-fifths of the world's population, has had the fastest

\(^{12}\) The speed of increase in the United Kingdom may have been as high or higher than the U.S. in the 1980s, but the U.K. started from a much lower level of inequality. Atkinson (1997) estimates that between 1978 and 1992, the Gini coefficient rose more than 10 points, bringing the U.K. from one of the lower inequality levels in Western Europe to the highest.
economic growth of any country from 1980 until 2009, which is probably the fastest sustained economic growth in history. China has also had the largest rise in inequality of any country in the sample, rising from a Gini coefficient of 22.4 in 1985 to 44.9 in 2003. This increase is 10 points higher than the next highest increase in the sample (Armenia) and more than twice as high as the third highest increase.

Is China's extraordinary rise in inequality the result of its rapid economic growth? This question is beyond the scope of this paper, but a good point of comparison is Vietnam. Since implementing market liberalization about 1990, Vietnam has been the fifth fastest-growing economy in the world. It is among the larger nations of the world, with the 13th largest population, and it emerged from centrally planning a decade after China, with a similar path of market-oriented reforms. Vietnamese inequality rose in the first decade after reforms, but inequality has not risen since. Although almost all countries transitioning from central planning initially see increases in inequality (as in Eastern Europe and the former Soviet Union), Vietnam's recent experience suggests that China's continuing rise in inequality is not the inevitable consequence of

13 This is ignoring Equatorial Guinea, a tiny country that discovered oil. The growth rates are calculated from the Penn World Tables 7.0 (Heston et al., 2011), using real GDP per capita (rgdpch). China's income level grew at 8.9% or 7.4% per year from 1980 to 2009 depending on which of the two Penn World Table estimates of China's GDP one uses.

14 Vietnam's GDP per capita grew at 5.9% from 1990 to 2009 in the manner explained in the previous footnote, behind Equatorial Guinea, China, Bosnia and Herzegovina, and Trinidad and Tobago.
rapid economic growth or market liberalization.

The rapid rise in inequality in recent decades has made the U.S. the
great outlier of developed countries and China the outlier among
developing countries. Their inequality rise is much faster than would be
predicted by convergence of low-inequality countries. There is no evidence
that within-country inequality is increasing in the world as a whole.

The convergence of inequality is consistent with two recent regional
trends. Latin American countries have some of the highest inequality
levels in the world, and have experienced some of the fastest declines in
inequality in the past couple of decades. The highest income countries,
which have some of the lowest inequality levels in the world, have
experienced increasing inequality in recent decades. Both these trends are
shown in Figure 11 and Table 7.

### Table 7: Regional fixed effects trend lines

<table>
<thead>
<tr>
<th></th>
<th>High GDP OECD</th>
<th>Latin America</th>
<th>E. Europe &amp; FSU</th>
<th>Emerging Asia</th>
<th>Low &amp; Mid GDP</th>
</tr>
</thead>
<tbody>
<tr>
<td>ln(GDP p.c.)</td>
<td>4.625</td>
<td>-6.661</td>
<td>-0.532</td>
<td>3.713</td>
<td>-0.309</td>
</tr>
<tr>
<td></td>
<td>(2.95)**</td>
<td>(2.11)*</td>
<td>(0.32)</td>
<td>(2.47)*</td>
<td>(0.28)</td>
</tr>
<tr>
<td>Constant</td>
<td>-19.759</td>
<td>108.042</td>
<td>37.503</td>
<td>8.575</td>
<td>41.790</td>
</tr>
<tr>
<td></td>
<td>(1.21)</td>
<td>(3.89)**</td>
<td>(2.49)*</td>
<td>(0.70)</td>
<td>(4.36)**</td>
</tr>
<tr>
<td>(R^2)</td>
<td>0.12</td>
<td>0.12</td>
<td>0.00</td>
<td>0.15</td>
<td>0.00</td>
</tr>
<tr>
<td>(N)</td>
<td>223</td>
<td>213</td>
<td>193</td>
<td>69</td>
<td>628</td>
</tr>
</tbody>
</table>

Robust \(t\) statistics in parentheses

a High GDP OECD includes OECD90 countries with GDP per capita above $25,000.
b Emerging Asia includes Asian countries with GDP per capita between $1,500 and $10,000.
c Low & Mid GDP are all countries with GDP per capita below $25,000.

Three other regions have a mixture of below- and above-average
inequality levels. Eastern Europe and the former Soviet Union countries as a whole show no significant trend inequality (since 1994 by which time inequality in many of the countries had risen significantly from levels under central planning). “Emerging Asia” is made up of Asian countries with GDP per capita between $1,500 and $10,000. Two thirds of countries in this group have rising inequality in the data, and the region as a whole has a significant rise in inequality. Finally, all the low- and middle-income countries (with GDP per capita below $25,000) show a slight, but insignificant, decline in inequality.

6 Conclusion

This study is the first to test for convergence of inequality using internally consistent time series of inequality for a large number of countries. The data show strong evidence of convergence. High inequality countries have seen their inequality levels fall, on average, as income rose in recent decades, and low inequality countries have seen their inequality levels rise on average. Using three different methods, linear trends, piecewise linear trends, and stochastic kernel estimation, the typical trend in less equal countries was the opposite of the trend in more equal countries, pushing inequality levels in both groups towards each other.

15 China, Indonesia, Iran, Sri Lanka, the Philippines, Thailand, Taiwan, and Vietnam all have rising inequality trends in the sample, and Jordan, Malaysia, Pakistan, and Turkey have falling inequality trends. Asia as a whole has a rising, but statistically insignificant trend.
The evidence for inequality convergence is consistent with the apparent narrowing of inequality as income rises in Figure 1. However, the dramatically falling trend in average inequality in Figure 1 is not so clear once one accounts for differences in country inequality levels. There is no clear trend of inequality within countries in the linear analysis. The stochastic kernel estimation shows lower income countries tending towards dramatically lower inequality levels in the stationary state, and higher income countries tending towards a modestly higher inequality. It is not clear why the linear and stochastic kernel results for overall inequality trends are different. The stochastic kernel estimation predicts that the Gini coefficient of countries with GDP per capita below $15,000 is predicted to fall from a current average of 41.3 to an average of 27.1 in the stationary state. The Gini coefficient of countries with GDP per capita above $15,000 is predicted to rise from an average of 29.5 currently to an average of 32.1 in the stationary state.

The patterns found in this new inequality data could be viewed good or bad news. The good news is that there is no sign of a tendency for low-income countries as a group to experience worsening inequality as their economy grows, unlike Kuznets's prediction. High inequality countries in particular, whether lower income or higher income, have tended to see substantial declines in inequality with economic growth. There also evidence using stochastic kernel estimation for a decline in inequality for all low income countries as they approach higher income
levels.

The bad news is that countries with unusually low inequality tend to see increases in inequality. The recent increases in inequality in higher income countries fit this trend of rising inequality.

One can speculate about possible reasons for a convergence of inequality with economic growth. First, trade and other forces of globalization may cause the income earned by different segments of society to become more similar across countries as income grows. The share of trade and other links to the outside world tend to grow with income levels. The convergence of inequality due to trade may be evidence of factor price equalization predicted by the Hecksher-Ohlin trade model (see Deardorff, 2001, for example).

Second, democratic participation tends to increase with income levels, and greater political activism by lower income citizens could change the income distribution through government tax rates and transfers, as well as union rights, government funding for education and health and other policies. Political pressure for policies that reduce income inequality is more likely to be strongest in countries that start out with higher inequality, leading to convergence.

A third possible cause of convergence comes from growth models with credit constraints, as noted by Bénabou in the introduction. The credit constraints cause inequality due to the asymmetric advantages in saving and investment of the initially unconstrained higher income
households. As income levels rise, fewer and fewer households remain constrained. The role of inequality in theoretical models of growth is surveyed in Bertola, et al. (2005).

A fourth dynamic by which economic growth could lead to inequality convergence is diversity in land ownership concentration across countries. Countries with more unequal ownership of agricultural land would have higher inequality of farm incomes. Since the share of agriculture diminishes in the economy as it grows, the role of land ownership in inequality would diminish, pushing countries towards more similar inequality levels. One would expect this dynamic to be most important at low income levels when agriculture is a large share of GDP, but in the data, inequality convergence is also clear in middle- and high-income countries as well.

A challenge for empirical research on inequality is that everything that influences incomes (employment, capital markets, education, trade, social welfare programs, savings, etc., etc.) also affects income distribution. This makes it difficult to isolate the specific causes of inequality convergence, just as empirical research on the causes of economic growth is difficult because almost everything affects GDP levels.

The convergence of inequality is in stark contrast to the general lack of convergence of income levels across countries. The convergence of inequality occurs when there is income growth, so if low-income countries fail to grow economically they would not be exposed to this tendency.
The convergence of inequality will do little to improve the world distribution of income if the poorest countries are growing the slowest. However, if a poor country with high inequality is able to grow economically, there is a good chance that inequality will also fall.
References


**Income Distribution** 10(1-2):5-12.


Figure 1: Inequality Versus Income Series for 87 Countries

Figure 2: Spuriously estimated inequality decline
Figure 3: Linear and quadratic fixed effect inequality trends

Figure 4: Linear and quadratic inequality trends on a logarithmic scale
Figure 5: Low, middle, and high income linear trends on a log scale

Figure 6: Bivariate kernel density estimates of transition matrix
Figure 7: Stationary kernel with low and high income inequality

Figure 8: Stationary kernels below and above $15,000 GDP p.c.
Figure 9: U.S. and high-income OECD inequality trends

Figure 10: China and low-income inequality trends
Figure 11: Regional fixed effects inequality trends

a “High GDP OECD” includes OECD90 countries with a GDP per capita above $25,000
b “Emerging Asia” includes Asian countries with GDP per capita levels between $1,500 and $10,000.
c Low & Mid GDP are all countries with GDP per capita below $25,000.
Online Data Appendix

This appendix lists the sources and definitions used for each of the countries in the sample.

Table A.1 shows the sources of inequality data along with the income concept used and the within-household weights applied.

<table>
<thead>
<tr>
<th>Source</th>
<th>No. of countries</th>
<th>Income concept</th>
<th>Household weights*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eurostat</td>
<td>21</td>
<td>Disposable Income</td>
<td>Equivalent 1</td>
</tr>
<tr>
<td>TransMONEE</td>
<td>19</td>
<td>Disposable Income</td>
<td>Per capita</td>
</tr>
<tr>
<td>SEDLAC</td>
<td>18</td>
<td>Disposable Income</td>
<td>Equivalent 2</td>
</tr>
<tr>
<td>LIS</td>
<td>4</td>
<td>Disposable Income</td>
<td>Per capita</td>
</tr>
<tr>
<td>WIID</td>
<td>22</td>
<td>Varied</td>
<td>Varied</td>
</tr>
<tr>
<td>Other</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Canada</td>
<td></td>
<td>Disposable Income</td>
<td>Equivalent 3</td>
</tr>
<tr>
<td>Indonesia</td>
<td></td>
<td>Consumption</td>
<td>Per capita</td>
</tr>
<tr>
<td>Vietnam</td>
<td></td>
<td>Consumption</td>
<td>Equivalent 1</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>87</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Equivalent 1 = Modified OECD adult equivalent weights; Equivalent 2 = Household equivalent D (from SEDLAC, 2011); Equivalent 3 = Square root of family size. See online data appendix for details.

Household equivalency scales

The Gini coefficients are calculated with equivalized household income obtained by dividing total household income by factors which differ according to the equivalency scale. Equivalency scales capture the sensible idea that there are economies of scale in household consumption, but there is no generally accepted equivalency scale, and Deaton (1992)
argues that it is not possible to estimate household equivalency scales because they inherently unidentifiable.

In the Table A.2, the equivalency scales depend on:
- \( H \) - household head (the first adult)
- \( A \) - number of other adults (including children aged 14 and over)
- \( K_1 \) - number of children aged 5 or younger
- \( K_2 \) - number of children between age 6 and 13 years (inclusive)

<table>
<thead>
<tr>
<th>Household equivalency scale</th>
<th>Weights</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Household</td>
<td>( H )</td>
</tr>
<tr>
<td>Household per capita</td>
<td>( H + A + K_1 + K_2 )</td>
</tr>
<tr>
<td>Household eq, OECDmod (OECD, 2005)</td>
<td>( H + 0.5A + 0.3(K_1 + K_2) )</td>
</tr>
<tr>
<td>Household eq, D (SEDLAC, 2010)</td>
<td>((H + A + 0.3K_1 + 0.5K_2)^{0.75})</td>
</tr>
<tr>
<td>Household eq, square root</td>
<td>((H + A + K_1 + K_2)^{0.5})</td>
</tr>
</tbody>
</table>

Household eq, D was chosen because it is the SEDLAC scale closest to the OECD modified scale for typical family sizes.

**Income Definitions**

Disposable Income  
after tax household income. There are some differences across countries in the measurement of in kind income, profits, rents, capital income, government transfers, implicit rent from home ownership, etc.

Gross Income  
Same as disposable income, except it includes taxes.

Consumption  
Household expenditure, generally excluding expenditure on durables.

**Sources**

Sources are citations to the references, except when they are prefixed by “WIID”, in which case they refer to the source referenced in WIID (2008).
<table>
<thead>
<tr>
<th>Country</th>
<th>Years</th>
<th>Equivalency scale</th>
<th>Income Definition</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Western Europe (Eurostat)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Croatia</td>
<td>2002-2008</td>
<td>Household eq. OECDmod</td>
<td>Disposable Income</td>
<td>Eurostat 2011</td>
</tr>
<tr>
<td>Denmark</td>
<td>95.97,99,01,03-09</td>
<td>Household eq. OECDmod</td>
<td>Disposable Income</td>
<td>WIID:European Commission; Eurostat 2011 (2001-2009)</td>
</tr>
<tr>
<td>Norway</td>
<td>2003-2009</td>
<td>Household eq. OECDmod</td>
<td>Disposable Income</td>
<td>Eurostat 2011</td>
</tr>
<tr>
<td>Turkey</td>
<td>2002-2003</td>
<td>Household eq. OECDmod</td>
<td>Disposable Income</td>
<td>Eurostat 2011</td>
</tr>
<tr>
<td>Eastern Europe and the former Soviet Union (TransMONEE)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Armenia</td>
<td>1996, 2002-2008</td>
<td>Household per capita</td>
<td>Disposable Income</td>
<td>Transmonee 2010</td>
</tr>
<tr>
<td>Bulgaria</td>
<td>1994-2008</td>
<td>Household per capita</td>
<td>Disposable Income</td>
<td>Transmonee 2010</td>
</tr>
<tr>
<td>Belarus</td>
<td>1995-2008</td>
<td>Household per capita</td>
<td>Disposable Income</td>
<td>Transmonee 2010</td>
</tr>
<tr>
<td>Czech Republic</td>
<td>1994-2007</td>
<td>Household per capita</td>
<td>Disposable Income</td>
<td>Transmonee 2010</td>
</tr>
<tr>
<td>Estonia</td>
<td>1995-2005</td>
<td>Household per capita</td>
<td>Disposable Income</td>
<td>Transmonee 2010</td>
</tr>
<tr>
<td>Hungary</td>
<td>1994-2007</td>
<td>Household per capita</td>
<td>Disposable Income</td>
<td>Transmonee 2010</td>
</tr>
<tr>
<td>Kyrgyz Republic</td>
<td>1997-2008</td>
<td>Household per capita</td>
<td>Disposable Income</td>
<td>Transmonee 2010</td>
</tr>
<tr>
<td>Latvia</td>
<td>1994-2004</td>
<td>Household per capita</td>
<td>Disposable Income</td>
<td>Transmonee 2010</td>
</tr>
<tr>
<td>Lithuania</td>
<td>1997-2004</td>
<td>Household per capita</td>
<td>Disposable Income</td>
<td>Transmonee 2010</td>
</tr>
<tr>
<td>Macedonia, FYR</td>
<td>1994-2008</td>
<td>Household per capita</td>
<td>Disposable Income</td>
<td>Transmonee 2010</td>
</tr>
<tr>
<td>Country</td>
<td>Years</td>
<td>Equivalency scale</td>
<td>Income Definition</td>
<td>Source</td>
</tr>
<tr>
<td>--------------</td>
<td>----------------------</td>
<td>----------------------------</td>
<td>--------------------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>Poland</td>
<td>1994-2008</td>
<td>Household per capita</td>
<td>Disposable Income</td>
<td>Transmonee 2010</td>
</tr>
<tr>
<td>Romania</td>
<td>1994-2008</td>
<td>Household per capita</td>
<td>Disposable Income</td>
<td>Transmonee 2010</td>
</tr>
<tr>
<td>Serbia</td>
<td>2003-2008</td>
<td>Household per capita</td>
<td>Disposable Income</td>
<td>Transmonee 2010</td>
</tr>
<tr>
<td>Slovak Republic</td>
<td>1996-2008</td>
<td>Household per capita</td>
<td>Disposable Income</td>
<td>Transmonee 2010</td>
</tr>
<tr>
<td>Latin America and Caribbean (SEDLAC)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Belize</td>
<td>93,94,97-99</td>
<td>Household eq, D</td>
<td>Gross Income</td>
<td>SEDLAC, 2011</td>
</tr>
<tr>
<td>Bolivia</td>
<td>97,99,01,02,05-07</td>
<td>Household eq, D</td>
<td>Gross Income</td>
<td>SEDLAC, 2011</td>
</tr>
<tr>
<td>Brazil</td>
<td>81-90,93,95-99,01-09</td>
<td>Household eq, D</td>
<td>Gross Income</td>
<td>SEDLAC, 2011</td>
</tr>
<tr>
<td>Chile</td>
<td>87,90,92,94,96,98,00,03,06,09</td>
<td>Household eq, D</td>
<td>Gross Income</td>
<td>SEDLAC, 2011</td>
</tr>
<tr>
<td>Colombia</td>
<td>96,99-01,03,04,06</td>
<td>Household eq, D</td>
<td>Gross Income</td>
<td>SEDLAC, 2011</td>
</tr>
<tr>
<td>Ecuador</td>
<td>95,98,99,03-09</td>
<td>Household eq, D</td>
<td>Gross Income</td>
<td>SEDLAC, 2011</td>
</tr>
<tr>
<td>Guatemala</td>
<td>2002-04</td>
<td>Household eq, D</td>
<td>Gross Income</td>
<td>SEDLAC, 2011</td>
</tr>
<tr>
<td>Jamaica</td>
<td>90-96,99,01-02</td>
<td>Household eq, D</td>
<td>Gross Income</td>
<td>SEDLAC, 2011</td>
</tr>
<tr>
<td>Mexico</td>
<td>89,92,94,96,98,00,02,04-06,08</td>
<td>Household eq, D</td>
<td>Gross Income</td>
<td>SEDLAC, 2011</td>
</tr>
<tr>
<td>Nicaragua</td>
<td>93,98,01,05</td>
<td>Household eq, D</td>
<td>Gross Income</td>
<td>SEDLAC, 2011</td>
</tr>
<tr>
<td>Panama</td>
<td>89,91,95,97-98,01-06,09</td>
<td>Household eq, D</td>
<td>Gross Income</td>
<td>SEDLAC, 2011</td>
</tr>
<tr>
<td>Paraguay</td>
<td>95,97,99,01-09</td>
<td>Household eq, D</td>
<td>Gross Income</td>
<td>SEDLAC, 2011</td>
</tr>
<tr>
<td>Uruguay</td>
<td>2006-09</td>
<td>Household eq, D</td>
<td>Gross Income</td>
<td>SEDLAC, 2011</td>
</tr>
<tr>
<td>Venezuela</td>
<td>89,92,95,98-2006</td>
<td>Household eq, D</td>
<td>Gross Income</td>
<td>SEDLAC, 2011</td>
</tr>
<tr>
<td>Luxembourg Income Study and others</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Switzerland</td>
<td>82,92,00,02,04</td>
<td>Household per capita</td>
<td>Disposable Income</td>
<td>LIS, 2011</td>
</tr>
<tr>
<td>United States</td>
<td>74, 79, 86, 91, 94, 97, 00, 04</td>
<td>Household per capita</td>
<td>Disposable Income</td>
<td>LIS, 2011</td>
</tr>
<tr>
<td>Israel</td>
<td>79, 86, 92, 97, 01, 05</td>
<td>Household per capita</td>
<td>Disposable Income</td>
<td>LIS, 2011</td>
</tr>
<tr>
<td>Taiwan</td>
<td>81, 86, 91, 95, 97, 00, 05</td>
<td>Household per capita</td>
<td>Disposable Income</td>
<td>LIS, 2011</td>
</tr>
<tr>
<td>Canada</td>
<td>1976-2004</td>
<td>Household eq, square root</td>
<td>Disposable Income</td>
<td>Heinz, Table A.4 (p. 55)</td>
</tr>
<tr>
<td>Indonesia</td>
<td>76,78,80,81,84,87,90,93,96,99,02-09</td>
<td>Household per capita</td>
<td>Consumption</td>
<td>WIID, Statistics Indonesia; Statistics Indonesia, 2011 (1999-2010)</td>
</tr>
<tr>
<td>Country</td>
<td>Years</td>
<td>Equivalency scale</td>
<td>Income Definition</td>
<td>Source</td>
</tr>
<tr>
<td>-------------</td>
<td>----------------------------</td>
<td>-------------------</td>
<td>-------------------</td>
<td>------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Vietnam</td>
<td>93, 98, 02, 04, 06, 08</td>
<td>Household eq, OECDmod</td>
<td>Consumption</td>
<td>author's calculations</td>
</tr>
<tr>
<td>Australia</td>
<td>81, 85, 89, 95-98, 00-03, 05</td>
<td>Household eq, OECDmod</td>
<td>Disposable Income</td>
<td>WIID: Australian Bureau of Statistics 2005; Australian Bureau Statistics, 2009 (Table S6, p. 11, for 2003 &amp; 2005)</td>
</tr>
<tr>
<td>Bahamas</td>
<td>73, 75, 77, 79, 86, 88, 89, 91-93, 03, 04</td>
<td>Total household</td>
<td>Gross Income</td>
<td>WIID: Bahamas Dept. of Statistics, Household Sample Survey</td>
</tr>
<tr>
<td>Bangladesh</td>
<td>83, 86, 88, 92, 96, 00, 05</td>
<td>Household per capita</td>
<td>Consumption</td>
<td>WIID: Bangladesh Household Expenditure Survey</td>
</tr>
<tr>
<td>China</td>
<td>85, 91, 95, 00, 03</td>
<td>Household per capita</td>
<td>Disposable Income</td>
<td>WIID: Chotikapanich et al 2005: Rural/Urban Household Survey</td>
</tr>
<tr>
<td>Lesotho</td>
<td>1986, 1993</td>
<td>Household per capita</td>
<td>Consumption</td>
<td>WIID: World Bank</td>
</tr>
<tr>
<td>New Zealand</td>
<td>86, 89, 92, 95, 98, 01, 04</td>
<td>Household eq, square root</td>
<td>Disposable Income</td>
<td>WIID: Perry 2005</td>
</tr>
<tr>
<td>Pakistan</td>
<td>87, 91, 93, 96, 02, 05</td>
<td>Household per capita</td>
<td>Consumption</td>
<td>WIID: World Bank</td>
</tr>
<tr>
<td>Philippines</td>
<td>85, 88, 91, 94, 97, 00, 03</td>
<td>Household per capita</td>
<td>Consumption</td>
<td>WIID: World Bank</td>
</tr>
<tr>
<td>Thailand</td>
<td>81, 86, 88, 90, 92, 94, 96, 98, 99</td>
<td>Household per capita</td>
<td>Consumption</td>
<td>WIID: Thailand Socio-Economic Survey</td>
</tr>
</tbody>
</table>
References


Figure A1: Year of inequality observations
Figure A2: Number of inequality observations per country
Urbanization without Structural Transformation: Evidence from Sub-Saharan Africa

Douglas GOLLIN\textsuperscript{a} Rémi JEDWAB\textsuperscript{b} Dietrich VOLLRATH\textsuperscript{c}

\textsuperscript{a} Department of International Development, University of Oxford
\textsuperscript{b} Department of Economics, George Washington University
\textsuperscript{c} Department of Economics, University of Houston

This Version: November 13th, 2012

Abstract: In most economies across space and time, urbanization has been accompanied by structural transformation. Standard theories of structural change understand the process of urbanization as driven to some degree by productivity growth in agriculture (Green Revolutions) or in non-agriculture (Industrial Revolutions). Present-day African economies appear to be anomalous with respect to this theoretical framework, however. Rapid urbanization has taken place across Africa in spite of very low growth in productivity in either agriculture or industry. This paper explores the forces driving African urbanization. We argue that a key part of the explanation is the reliance of many African countries on exports of natural resources. With non-homothetic preferences, the ensuing resource rents are disproportionately spent on urban goods and services. This drives urbanization through the rise of “consumption cities”. By contrast, the urbanization process in much of Asia took place through the rise of “production cities” that export manufactured goods to the rest of the world. We develop a new model of structural transformation and trade that investigates the different development trajectories of economies that have comparative advantages in tradable manufactured goods or, alternatively, natural resource exports. We argue that this model may help to explain Africa’s urbanization without development.

Keywords: Urbanization; Structural Change; Resource Curse; Africa

JEL classification: L16; N17; O18; O40; O55; R10
1. INTRODUCTION

In 1950, most people across the developing world lived in rural areas and worked in agriculture. The urbanization rate in both Africa and developing Asia was 10-15%, approximately the same level as in Renaissance Europe. Most of the countries of the developing world were also poor, and widespread poverty persisted for several decades. As late as 1981, in the countries defined today by the World Bank as “low-and middle-income economies”, 70% of the population fell below the $2/day poverty line.

In the years since 1950, urbanization has proceeded at rapid rates across the developing world. In both Asia and Africa urban shares of the population reached around 40% by 2010 – roughly comparable to the level in today’s rich countries in the years following the Industrial Revolution. In the developing countries of Asia, urbanization has been accompanied by substantial economic growth, as happened previously in Europe and North America.

Somewhat curiously, however, Africa has urbanized with relatively little growth. As Figure 1 shows, urbanization in Africa has occurred at approximately the same pace as urbanization in Asia. But Africa remains far poorer, by almost any measure, than developing Asia.

Why is this a puzzle? Urbanization is usually seen as a consequence of economic development. As a country develops, the process of structural transformation implies that people move out of rural areas and agricultural employment. They move into urban-based manufacturing and service sectors. In a closed-economy model, a subsistence requirement for food consumption almost mechanically implies that development parallels urbanization. With standard Engel curves, economic growth implies increasing consumption of non-agricultural goods and services that are typically produced in urban areas. While this model seems to hold reasonably well for Asia, Africa’s urban growth took place in a period with only modest economic growth and without widespread industrialization. So why did Africa urbanize?

Asia is a perfect example of urbanization with structural transformation. Standard structural transformation models distinguish labor push and labor pull factors as the main drivers of the rural-urban transition (Alvarez-Cuadrado & Poschke, 2011). The labor push approach shows how a rise in agricultural productivity - a Green Revolution - reduces the “food problem” and releases labor for the modern sector (Schultz, 1953; Gollin, Parente & Rogerson, 2002, 2007; Michaels, Rauch & Redding, 2012). The labor pull approach describes how a rise in non-agricultural productivity - an Industrial Revolution - attracts underemployed labor from agriculture into the modern sector (Lewis, 1954; Harris & Todaro, 1970; Hansen & Prescott, 2002; Lucas, 2004). Alternatively, a country with a comparative advantage in manufacturing can open up to trade and use imports to solve its food problem (Matsuyama, 1992; Teigner, 2011; Yi & Zhang, 2011). These mechanisms all lead to greater non-agricultural employment, and thus greater urban employment as a proportion of total workforce. The successful Asian economies all went through both a Green Revolution and an Industrial Revolution, with development following corresponding patterns (Evenson & Gollin, 2003; Young, 2003; Bosworth & Collins, 2008; Brandt, Hsieh & Zhu, 2008; McMillan & Rodrik, 2011).

In contrast, Africa offers a perfect example of urbanization without structural transformation. First, there has been little evidence of a Green Revolution in Africa. Its food yields have remained low (Evenson & Gollin, 2003; Caselli, 2005; Restuccia, Yang & Zhu, 2008); in 2009, cereal yields were 2.8 times lower than in Asia, while yields were 2.1 times lower for starchy roots. Second, there has been no Industrial Revolution in Africa. Its manufacturing and service sectors are relatively small and unproductive (McMillan & Rodrik, 2011; Badiane, 2011); in 2007, employment shares in industry and services were 10% and 26% for Africa, but 24% and 35% for Asia, and African labor productivity was 1.7 and 3.5 times lower in industry and services, respectively (World Bank, 2010).

We argue that part of Africa’s differential pattern of urbanization can be explained by its dependence on natural resource exports, associated with what can be described as a Natural Resource Revolution (Jedwab, 2011). Africa is richly endowed with mineral and fuel reserves. In addition, many African countries appear to have a strong comparative advantage in the production and export of various non-food crops, such as cocoa, coffee, tea, cotton, rubber, or timber (Tosh, 1980). We view these production activities as a form of...
natural resource extraction, in the sense that these crops take advantage of natural resource availability and produce goods that have essentially no domestic market.

The close relationship between Africa’s urbanization and its natural resource can be seen in Figure 2, which plots the urbanization rate against GDP shares of “manufacturing and services” and “natural resource exports” (fuels, minerals and cash crops) for Asia and Africa in 2000. While Asian urbanization is associated with manufacturing and services, the most urbanized African countries are heavily dependent on exports of natural resources.1

The effect of resource exports on urbanization could occur through a number of different channels, which we discuss in the African context. One possible channel is that the production or marketing of natural resources could be concentrated in cities. By itself, however, this would not seem to explain the link between natural resource extraction and urbanization. Point-source natural resources (e.g., oil and minerals) are highly capital-intensive, and production of these commodities creates very little direct employment.2 Cash crops and timber are produced in rural areas and contribute to rural employment, rather than urban employment.

A second possible link is that resource exports could have backward and forward production linkages with the urban-based sectors. For example, input marketing and/or output processing could generate jobs in urban areas. In practice, however, mining equipment and inputs are typically imported rather than produced domestically, and the mode of production of cash crops has remained traditional in Africa, with limited use of modern inputs. And there is little or no local processing or transformation of natural resource outputs in Africa; famously, Africa imports its own coffee and tea after they have been processed in Europe.3

We focus instead on a third channel of causation. We argue that differences between international prices and local production costs imply that natural resource exports have generated considerable surplus for African countries. If the Engel curve – non-homothetic preferences – implies that this surplus is mainly spent on urban goods and services, this drives urbanization through the rise of consumption cities. Because greater opportunities in the urban sector attract people to cities, the model is in line with the labor pull hypothesis, except it considers trade and the resource sector as the main drivers of this transformation. As wealth is created in the rural cash crop sector or urban mining sector, it is spent in the urban non-tradable sector, with demand for these goods driving urbanization. These consumption cities differ from the Asian production cities where wealth originates from urban-based production activities such as tradable manufacturing or services.4

We develop a model of structural transformation and trade that captures the three types of urbanization. In the model economy, urbanization can be driven by a Green Revolution (an increase in agricultural productivity), an Industrial Revolution (an increase in manufacturing or service productivity), or a Resource Revolution (which can be thought of as being driven by high productivity or a high price in the resource sector).

The model considers two small open economies that will differ in their resource endowments. Each economy has two factors of production – unskilled labor and skilled labor. We consider four distinct production sectors: food, natural resources, and other tradable and non-tradable goods and services. We

---

1In 2000, 27 out of 41 Sub-Saharan African countries and 4 out of 22 Asian countries have more than 10% of their GDP coming from natural resource exports. In 1960, there were 31 African countries and 3 Asian countries satisfying this criterion. We obtain an Asian pattern for Latin American countries and an African pattern for Middle East and North African countries.

2For instance, Angola’s urbanization rate was 15% before oil was discovered there, and it is 60% in 2010. Crude oil now accounts for over 50% of GDP but it employs fewer than 10,000 nationals. Botswana has a similar urbanization rate, but while the diamond sector accounts for 36% of GDP it provides employment for approximately 13,000 people.

3The reasons for the lack of local transformation of natural resources in Africa (e.g., oil refining or cash crop processing) is beyond the scope of this paper. Transformation is capital-intensive, and downstream marketing may require specialized knowledge of consumer tastes. Poor physical infrastructure and weak institutions may have reduced the incentives for firms to locate processing industries in the developing world, but this is largely speculation. For whatever reasons, however, African economies perform very little processing or transformation of natural resource raw materials.

4Surely, there are consumption cities in Asia, as in Mongolia, and production cities in Africa, as in South Africa and Botswana. But overall we believe that the patterns found in Figure 2 show how distinct Asia and Africa are.
think of tradables and non-tradables as urban-based activities corresponding to the non-agricultural goods that people favor after they have satisfied their subsistence consumption needs. In addition to services such as haircuts, bars and restaurants, construction, health care, and retailing, we think of the non-tradable sector as producing some manufactured goods that are not traded – such as the local production of some goods such as clothing, footwear, furniture, and customized metal work (e.g., metal gates and window bars for houses). Any relative expansion of these two sectors will raise the urbanization rate.

One of the economies has a factor endowment that gives it a comparative advantage in natural resources. We think of this as our archetypal African economy. The other economy, which we think of as a representative Asian country, has a comparative advantage in tradables. We abstract from questions about the sources of these comparative advantages. There is a lengthy literature arguing that Africa’s dependence on natural resource exports is the end result of weak institutions, rather than some intrinsic characteristic based on geographic determinism. We accept this point of view. Our purpose is not to focus on the sources of comparative advantage or to untangle the relative importance of institutional factors and geographic factors as causes of development. Instead, we ask how this pattern, once in place, may drive different types of urbanization.

In autarky, both of our model economies are poor and rural. In either economy, a Green Revolution will have the effect of alleviating the food problem and releasing labor for other sectors. This allows the countries to specialize in their sectors of comparative advantage. Our Asian country exports tradables to the rest of the world. As the production of tradables takes place in cities, this country urbanizes through the rise of production cities. The African country exports natural resources to the rest of the world. As the Engel curve implies that the surplus of the resource sector is spent in the cities, this country also urbanizes – but through the rise of consumption cities that produce non-tradables. We assume the production of tradables requires unskilled labor and skilled labor, while the production of resources and non-tradables uses unskilled labor.

Urbanization is thus associated with economic development – rising wages – in the long run in the Asian country. In contrast, our model African country remains relatively poorer for the same level of urbanization. Although very simple, our model accounts for the fact Africa is urbanized but poor. Figure 3 confirms that resource-rich countries are much more urbanized than resource-poor countries in Africa, although the difference in per capita income is not staggering.

This paper is related to a large body of work on the role of sectoral labor productivity in driving structural change; i.e. the decline in agriculture, the rise and fall of manufacturing, and the rise of services (see Herrendorf, Rogerson & Valentinyi 2011 for a survey of the literature). A first strand of the literature looks at the origins of structural change in developed countries. The labor push approach shows how a rise in agricultural productivity – a Green Revolution – reduces the “food problem” and releases labor for the modern sector (Schultz, 1953; Gollin, Parente & Rogerson, 2002, 2007; Nunn & Qian, 2011; Michaels, Rauch & Redding, 2012). The labor pull approach describes how a rise in non-agricultural productivity - an industrial revolution - attracts underemployed labor from agriculture into the modern sector (Lewis, 1954; Harris & Todaro, 1970; Hansen & Prescott, 2002; Lucas, 2004; Alvarez-Cuadrado & Poschke, 2011). A second strand of the literature studies whether income effects or price effects explain structural change. Non-homothetic preferences and rising incomes mean a reallocation of expenditure shares towards non-agricultural goods (Caselli & Coleman II, 2001; Gollin, Parente & Rogerson, 2002, 2007; Matsuyama, 1992, 2002; Voigtländer & Voth, 2006; Galor & Mountford, 2008; Duarte & Restuccia, 2010; Alvarez-Cuadrado & Poschke, 2011). Ngai & Pissarides (2007) and Acemoglu & Guerrieri (2008) see structural change as a consequence of price effects: assuming a low elasticity of substitution across consumption goods, any relative increase in the productivity of one sector leads to a relative decrease in its employment share. Buera & Kaboski (2009), Yi & Zhang (2011) and Michaels, Rauch & Redding (2012) adopt or compare both approaches. According to the income effects approach, any rise in sectoral productivity leads to a reallocation of labor from inferior goods to superior goods. According to the price effects approach, the patterns can only be explained by a rise in agricultural productivity followed by a rise in manufacturing productivity.

We make three contributions to the literature. First, we show how adding the resource sector to the
three-sector growth model with non-homothetic preferences produces urbanization. This explains how a country can urbanize without any change in agricultural and manufacturing productivity, particularly in an open economy setting (Corden & Neary, 1982; Matsuyama, 1992; Echevarria, 2008; Galor & Mountford, 2008; Teigner, 2011; Yi & Zhang, 2011). Price effects do not necessarily hold here, as any rise in sectoral productivity could increase employment if it leads to more exports. In our model, the African country has a comparative advantage in natural resources and urbanizes as a result of a strong income effect. The Asian country has a comparative advantage in tradable manufactured goods and services and urbanizes as a result of a strong specialization effect. Third, our paper offers a single framework that can account for the differential paths of urbanization in Africa and Asia. Our model gives an explanation for the fact that Africa is urbanized but poor and also contributes to an understanding of the fact that Africa’s decline in agricultural employment is associated with a rise in services but not in manufacturing.

Our second contribution relates to the literature on urbanization in developing countries. While the structural transformation literature cited above portrays urbanization as a by-product of economic development, the economic geography literature suggests that agglomeration promotes growth, in both developed countries (Rosenthal & Strange, 2004; Henderson, 2005) and developing countries (Overman & Venables, 2005; Henderson, 2010; Felkner & Townsend, 2011). Given that urbanization is a form of agglomeration, it has been argued that cities could promote growth in developing countries (Duranton, 2008; Venables, 2010; World Bank, 2009; McKinsey, 2011). The structural transformation and agglomeration economies could explain why urbanization appears as strongly correlated with economic development (Acemoglu, Johnson & Robinson, 2002; Henderson, 2010). Yet an optimistic view of urbanization in developing countries is perhaps difficult to square with empirical evidence on Africa. A few studies argue that Africa has urbanized without it being fully explained by economic development (Bairoch, 1988; Fay & Opal, 2000). This excessive urbanization is attributed to pull and push factors feeding rural exodus. Some argue that Africa’s urban growth can be attributed to rural poverty (Barrios, Bertinelli & Strobl, 2006; Poelhekke, 2010; de Janvry & Sadoulet, 2010). But the fact that urban growth was concentrated in resource-rich countries suggests that natural resources have been a more important factor. Others have focused on theories of urban bias, arguing that urban-biased policies have led to overurbanization and primacy in poor countries (Lipton, 1977; Bates, 1981; Ades & Glaeser, 1995; Davis & Henderson, 2003). However, there cannot be any urban bias without a surplus in the economy for governments to redistribute. In our view, resource-rich countries urbanize because of the surplus generated by resource exports; governments just amplify this effect by reallocating an even greater share of the surplus to cities. In our model, the African country is urbanized but poor because it urbanizes without tradables – tradable manufactured goods and services – and without human capital accumulation. Our paper thus provides an explanation for the growing “urbanization of global poverty” (Ravallion, Chen & Sangraula, 2007).

Finally, this paper contributes to the literature on Dutch disease and the resource curse (Corden & Neary, 1982; Matsuyama, 1992; Sachs & Warner, 2001; Robinson, Torvik & Verdier, 2006; Angrist & Kugler, 2008; Michaels, 2011; Caselli & Michaels, 2012). Dutch disease models explain why a country deindustrializes when its resource sector booms. The boom shifts labor away from the manufacturing sector into the non-tradable service sector, yet the net effect on urbanization is ambiguous. These models do not explain why African countries never industrialized and directly transitioned from agricultural economies into service economies. In our model, resource booms unambiguously increase urbanization through the income effect, and resource-rich countries fail to develop their manufacturing sectors because the country imports manufactured goods from the rest of the world. In some sense, we highlight a new dimension of the resource curse – the fact that resource booms drive urbanization through the rise of (relatively unproductive) consumption cities. In our model, the African country is relatively urbanized for its income level and does not accumulate human capital.

The paper is organized as follows: The next section describes in greater detail the differential patterns of structural transformation in Asia and Africa and provides a more detailed motivation for examining the
differences between the two. Section 3 outlines a model of structural transformation in a closed economy. Section 4 extends this model to an open economy. Section 5 examines the dynamics of the structural transformation process. Section 6 concludes.

2. DIFFERENTIAL PATTERNS OF URBANIZATION: ASIA AND AFRICA

This section continues the discussion of urbanization patterns developed in the introduction and provides additional evidence. We begin by carefully documenting a series of stylized facts regarding urbanization in Asia and Africa. Our data are drawn from a variety of sources and cover a period of time during which economies in both regions have undergone substantial urbanization.

2.1 Data Sources

The United Nations report the official urbanization rate for most countries by decade from 1960 to 2010 (http://esa.un.org/unpd/wup/index.htm), which we use for an initial examination of the data. A potential problem with comparing countries using these data is that different countries use different definitions for urbanization. Although our descriptive statistics rely on these data, we have begun to test the validity of the comparisons. For instance, the UN data include text explaining the definition of urbanization adopted by each country for each census year. Preliminary evidence indicates that most countries use a population threshold to identify cities; i.e. a locality is defined as "urban" if it has more than a specified number of inhabitants. The average threshold value is approximately the same across the set of African and Asian countries for which we have data (around 4,200 inhabitants). In future regression analysis, we will control directly for this threshold value.

Data on natural resource exports are more problematic. Commonly used data from the World Bank and COMTRADE are very imperfect because of omissions and because of problems related to the unit values of trade. We prefer to use data coming from the United States Geological Service, which has published annual country reports from 1960 to the present that have detailed and consistent estimates of the export share of mining and fuels. For data on agricultural exports, which we also view as a form of natural resource exports, we have relied on the UN Food and Agriculture Organization's online FAOSTAT database. Together with standard data sources on national income accounts and PPP-adjusted measures of income per capita, these sources will allow us to calculate most of the ratios that our analysis requires.

2.2 Urbanization in Africa and Asia

At the beginning of the 20th century, both Africa and Asia were poor and heavily rural. Their economies were heavily dependent on low productivity agriculture, and from the admittedly sparse data that are available, it appears that urbanization rates in both regions were less than 5% (Bairoch 1988), the same level as in Medieval Europe. Half a century later, the two regions still looked quite similar: poor and rural. By 1950 their urbanization rates had crept up to 10-15%, which was still as low as in Renaissance Europe. By the start of the 21st century, urbanization has now reached around 40% in both Africa and Asia, which is comparable to the level attained in today’s developed countries after the Industrial Revolution.

Urban growth has taken place in cities of all sizes. Although the literature emphasizes the growth of the largest cities in the developing world, urban growth has in fact taken place in cities of all sizes - large, medium, and small. The distribution of city size across regions is quite similar, too. For example, in 2010, there were 257 Asian and 60 African “mega-cities” with over 750,000 inhabitants. Since Asia is roughly four times more populous than Africa, this means that Africa and Asia have approximately the same number of megacities per capita. The megacities represent around 40% of the urban population in both continents.

But although urbanization patterns in Africa and Asia look quite similar for the past half century, there are two striking differences. The first is that Asia’s urbanization has occurred in conjunction with rapid growth
in per capita income. And the second is that urbanization in Asia has coincided with a rapid movement of people out of agriculture into manufacturing and industry, whereas in Africa urbanization has been accompanied primarily by a movement into services. Figure 1 shows a comparison of income growth and urbanization in Asia and Africa; it is clear that although African GDP per capita has lagged far behind that of Asia, urbanization has proceeded in more or less the same fashion in both regions.

### 2.3 Structural Transformation in Africa and Asia

African countries have always been highly specialized in the export of natural resources, whether mineral products, fuels or cash crops. In 1960, 31 out of 41 Sub-Saharan African countries and only 3 out of 22 Asian countries had more than 10% of their GDP coming from natural resource exports. In contrast, Asian countries developed through expansion of their manufacturing sectors, with industrialization successively sweeping through Japan, South Korea, Hong Kong, Singapore, Taiwan, Malaysia, Thailand and China.

These differences have persisted in the recent period. In 2000, 27 out of 41 Sub-Saharan African countries and 4 out of 22 Asian countries have more than 10% of their GDP coming from natural resource exports. By comparison, manufacturing and services account for 85.1% of GDP on average in Asia, against 54.4% in Africa.

While Asian urbanization is associated with manufacturing and services, the most urbanized African countries export natural resources. This is confirmed by Figure 2 which plots the urbanization rate against GDP shares of “manufacturing and services” and “natural resource exports” for Asia and Africa in 2000. We obtain an “Asian” pattern for Latin American countries and an “African” pattern for Middle East and North African countries.

Figure 1 shows that both Africa and Asia have followed the same urbanization patterns. Yet Africa’s urban growth was achieved in a period with only modest economic growth and without industrialization. Until the mid-1960s, Africa and Asia had the same level of per capita income. Asia has experienced strong economic growth in the subsequent period, as a result of industrialization and the rise of services. Resource exports have fostered African economic growth till the mid-1970s, but the subsequent period was characterized by macroeconomic disequilibria, social unrest and general impoverishment. The 1980s was a “lost decade” for Africa. Africa now has only one-fourth the income per capita of Asia, although it is equally urbanized. We will use the shorthand terminology of saying that Africa is “urbanized but poor”.

### 2.4 Different Paths among African Countries

Arguably there are many respects in which Africa differs from Asia, so it may not be convincing to seize on natural resource exports as the source of their different paths. However, we can also look within the African continent to compare countries with different levels of dependence on natural resource exports. It is hardly news to find a negative relationship between natural resource exports and income levels; this is simply the "resource curse" that has received much attention. But we note an additional fact: natural resource dependence within Africa seems to be closely (and positively) related to urbanization, which is not a point that is typically noted in the literature on resource curses.

Figure 3 shows the evolution of urbanization and per capita GDP for four groups of African countries depending on the average share of natural resource exports in GDP in 1960-2000. Urbanization rates were similar across the four country groups at independence, but resource rich countries are now much more urbanized (while the difference in per capita income is not staggering). For example, the urbanization rate of the resource-rich countries is 50.3% on average in 2010, while it is 25.1% on average for the resource-poor countries.

Moreover, when we look within countries, making use of time series data, it appears that resource-rich countries have all experienced dramatic urbanization after resource production boomed, whether as a
result of the discovery of mining deposits and oil fields or the adoption of agricultural innovations in the cash crop sector. Figure 4 shows urbanization patterns for eight African countries that have experienced a resource boom in the post-1960 period. In Gabon, Nigeria and Angola, urbanization boomed after oil was discovered in the early 1970s. In Liberia, urbanization increased after iron ore, gold and diamonds boomed in the 1960s-1970s. In Ghana and Ivory Coast, urbanization was largely driven by the cocoa booms, in the 1960s and 1990s in the former and in the 1960s, 1970s and 1980s in Ivory Coast (Jedwab 2012). In Zambia, copper boomed in the 1960s-1970s. Interestingly, urbanization decelerated after 1980 due to a collapse in copper prices. In Botswana, diamonds were discovered after 1973, and urbanization boomed as a result.

We can also look at those countries that were relatively urbanized at independence (taking as a threshold a 20% level of urbanization in 1960). In almost all cases, these were countries that had experienced resource booms in the pre-independence period. These were mostly countries that had experienced booms in agricultural commodities: rubber production boomed as early as the 1920s in the DRC, groundnut production boomed in the 1930s in Senegal, and cocoa production boomed in the 1930s in Ghana. In terms of mining exports, gold and diamonds were discovered in South Africa in the 1910s, while diamonds were discovered in the Central African Republic in 1930s.

By comparison, some Africa countries are still not very urbanized; these are countries with low exports of cash crops and little natural resource production. Figure 5 shows urbanization for 8 of these countries.

2.5 Central Questions

This cursory look at the data raises a series of questions that we will take up theoretically and with empirical analysis. The first question is whether these relationships have a plausible theoretical foundation and whether the data can be rationalized by a standard model of structural transformation. A second question is quantitative: how much urbanization could we have expected to see in Africa if it had been able to develop comparative advantages in sectors other than natural resource extraction. It is clear that urbanization would have increased anyway, as a result of demographic growth and the availability of food imports. Yet we believe that Africa's urbanization rates today are driven to high levels by natural resource exports that promote the development of “consumption cities” where natural resource rents are used to purchase non-tradable services. We are not convinced that these consumption cities display the same agglomeration externalities in production that have characterized urbanization in other parts of the world, which might account for the different relationship between urbanization and income levels that we currently observe.

3. A MODEL OF STRUCTURAL CHANGE

The evidence shows that structural change and urbanization are not monotonically related to income per capita. Africa from 1950–2010 experienced structural change and urbanization similar in magnitudes to that in East Asia, but without the rapid rise in living standards. Here we present a model that captures this distinction between the two regions. The model expands on existing models of structural change by allowing for several types of non-agricultural goods: tradable goods, non-tradable goods, and natural resources. Furthermore, we distinguish urban sectors (tradable and non-tradables) from rural sectors (resources and agriculture). These finer distinctions allow us to describe the processes of structural change and urbanization in more detail, which in turn provides a means to explain the distinction between Africa and Asia in the post-war period.

The driving difference between Africa and Asia is their comparative advantage in the production of natural resources versus tradable goods. The natural wealth of Africa led that continent to specialize in resource extraction, using the proceeds to import desired tradable goods. The ability to export resources raised incomes, leading to increased demand for other non-agricultural goods, in particular non-tradable goods. The increased demand for non-tradables leads to a structural shift towards that sector, which also involves
greater urbanization. However, non-tradable goods have relatively low human capital requirements and so Africa does not have the incentive to develop a skilled workforce. In contrast, Asia specializes in tradable goods, and this along with the associated income effect on non-tradables raises urbanization there. However, as a producer of tradable goods with relatively high human capital requirements, Asia faces incentives favoring more skilled (and hence higher paid) workers.

We show that even starting from essentially equal conditions around 1950, the difference in comparative advantage in the two regions is sufficient to introduce a gap in living standards between them despite experiencing similar shifts of workers into urban areas.

To proceed we first describe the a simple static model of the allocation of labor across different sectors. The model provides a richer description of urbanization than is present in most models of structural change, and allows for differences in the kinds of workers that locate in urban areas. Using that model, we examine the response to free trade in economies that have different comparative advantages. Specifically, we compare an economy with an advantage in producing natural resources to one with an advantage in producing tradable consumption goods. We show how even though both types of economies benefit from trade and begin urbanizing, they differ in the kinds of work being done in urban areas and hence in output per capita.

The section following this one expands on the static model to add in dynamics regarding choices the skill mix of the labor force and allows for technological change over time.

### 3.1 Individual Utility and Budgets

Individuals have a constant-elasticity of substitution utility function

\[
U_t = \left[ \theta_f^{1/\phi} (c_f - \bar{c}_f)^{(\phi-1)/\phi} + \theta_C^{1/\phi} C^{(\phi-1)/\phi} \right]^{\phi/(\phi-1)}
\]

where \(c_f\) is the amount of food consumed and \(\bar{c}_f\) captures a subsistence amount of food that all individuals must consume. This utility for the food good is typical in models of structural change, as it leads to an income elasticity less than one for food, consistent with Engel’s law.

The term \(C\) is an aggregator of non-food consumption items. The parameter \(\phi\) determines the degree to which individuals are willing to substitute between food and the non-food consumption goods. We will go forward assuming that \(\phi > 1\), implying that these goods are substitutes. This is consistent with explanations of structural change that emphasize changing relative prices, as opposed to income effects. A change in relative prices that makes \(C\) cheaper will, even holding income constant, lead to a shift out of food consumption and towards \(C\).

The budget constraint is

\[
p_f c_f + p_r c_r + p_n c_n + p_d c_d = w
\]

where \(p_j\) are the prices of the goods and \(w\) is the wage. Individuals earn only labor income. It is useful to re-write the budget constraint as follows

\[
p_f (c_f - \bar{c}_f) + p_r c_r + p_n c_n + p_d c_d = w - p_f \bar{c}_f
\]
where $w - p_f \bar{c}_f$ is the surplus income available to individuals after they have met their subsistence requirement for food.

Maximizing utility subject to this budget constraint, it is easiest to express the resulting demand functions in terms of expenditure shares. To aid in the exposition, define the following two price indices

$$P_C = \left[ p^1_{\theta_r} + p^1_{\theta_n} + p^1_{\theta_d} \right]^{1/(1-\epsilon)}$$

$$P = \left[ p^1_{\theta_f} + p^1_{\theta_c} \right]^{1/(1-\phi)}.$$  \hspace{1cm} (5)  \hspace{1cm} (6)

The first, $P_C$, is a price index of the non-food goods, while the second is the aggregate price index.

The fraction of surplus income that will be spent on net food consumption, $c_f - \bar{c}_f$ is denoted $\Omega_f$,

$$\Omega_f = \frac{p^1_{\theta_f}}{p^1_{\phi}},$$

and so total expenditures on net food are given by

$$P_f(c_f - \bar{c}_f) = (w - p_f \bar{c}_f) \Omega_f.$$  \hspace{1cm} (7)  \hspace{1cm} (8)

Given that the fraction $\Omega_f$ of surplus income is spent on food, the remaining fraction $1 - \Omega_f$ of surplus income is spent on the three non-food goods. The split of expenditures among the three goods is dictated by their price relative to the aggregate price $P_C$. Specifically, the fractions of surplus income expended good $j \in \{r, n, d\}$, denoted $\Omega_j$, is

$$\Omega_j = (1 - \Omega_f) \frac{p^1_{\theta_j}}{P^1_{\theta_C}}.$$  \hspace{1cm} (9)

It is tedious but straightforward to confirm that $\Omega_f + \Omega_r + \Omega_n + \Omega_d = 1$.

Once we have the production side of the economy in place, we will be able to use this demand structure to determine the allocation of labor across the various sectors of the economy.

3.2 Production and Factor Payments

The four goods are produced by technologies linear in labor. So for any given sector $j$ the production technology is

$$Y_j = A_j L_j$$

for $j \in \{f, r, n, d\}$, where $A_j$ is productivity in that sector and $L_j$ is the labor employed in that sector.

Labor is assumed to move freely between the sectors so that in autarky the wage is equivalent to

$$w = p_f A_f = p_r A_r = p_n A_n = p_d A_d$$

which will be used to determine the relative prices between various goods. There is an adding up constraint for labor employed across the sectors

$$L = L_f + L_r + L_d + L_n + L_d.$$  \hspace{1cm} (10)  \hspace{1cm} (11)  \hspace{1cm} (12)

To proceed, it will be useful to define two indices of productivity.

$$A_C = \left[ A^1_{\theta_r} + A^1_{\theta_n} + A^1_{\theta_d} \right]^{1/(\epsilon-1)}$$

$$A = \left[ A^1_{\theta_f} + A^1_{\theta_c} \right]^{1/(\phi-1)}.$$  \hspace{1cm} (13)  \hspace{1cm} (14)
Note that an increase in the productivity in any individual sector raises both indices. Furthermore, using the definition of the aggregate price index given previously, it is the case that the real wage is

$$\frac{w}{p} = A.$$  \hspace{1cm} (15)

Hence an increase in productivity in any sector raises the real wage.

### 3.3 Equilibrium in Autarky

Without trade, all sectors operate to provide goods to consumers. To begin solving for the equilibrium allocation of labor to the various sectors, note that (11) shows that the relative price between any two sectors can be written as $p_i/p_j = A_j/A_i$. In addition, the relative price of food to the non-food aggregate can be expressed as $p_f/p_c = A_c/A_f$.

Using these relative prices, we can re-write the fractions of expenditures in terms of productivity levels. We have that

$$\Omega_f = \frac{A^\phi - 1}{A^\phi - 1}.$$ \hspace{1cm} (16)

From this, and given our assumption that $\phi > 1$, an increase in the productivity of the food sector will raise the fraction of surplus income spent on food. On the other hand, an increase in the productivity of any of the non-food sectors will raise $A_j$ and lower the fraction spent on food as individuals substitute away to the cheaper non-food items.

Similarly, we can write the expenditure shares of the non-food items in terms of productivity levels,

$$\Omega_j = (1 - \Omega_f) \frac{A^\epsilon - 1}{A_C^\epsilon - 1}.$$ \hspace{1cm} (17)

for $j \in (r, n, d)$. Here, an increase in the productivity of sector $j$ has two conflicting effects. First, if $A_j$ rises, this lowers $\Omega_f$ as noted above, and this tends to increase $\Omega_j$. Secondly, though, an increase in $A_j$ will lower $\Omega_j$ due to our assumption that $\epsilon < 1$. An increase in the productivity of sector $j$ lowers the price of good $j$. But as complements, the lower price of good $j$ allows individuals to put more of their expenditures into other sectors.

To solve for the equilibrium allocation of labor across sectors, we equate total expenditures on a given sector to the total value of goods produced in that sector. For $j \in (r, n, d)$ we have that

$$p_j Y_j = L(w - p_f \bar{e}_f)\Omega_j,$$ \hspace{1cm} (18)

while for the food sector we have that

$$p_f Y_f = L(w - p_f \bar{e}_f)\Omega_f + Lp_f \bar{e}_f.$$ \hspace{1cm} (19)

Given the definition of the production functions, the free mobility of labor, and the definitions of the $\Omega$ terms, the solution for the fraction of labor employed in any given sector $j \in (r, n, d)$ is

$$\frac{L_j}{L} = \left(1 - \frac{\bar{e}_j}{A_j}\right) \Omega_j.$$ \hspace{1cm} (20)

while the fraction employed in food production is

$$\frac{L_f}{L} = \left(1 - \frac{\bar{e}_j}{A_j}\right) \Omega_f + \frac{\bar{e}_f}{A_f}.$$ \hspace{1cm} (21)

In each sector, the term $(1 - \bar{e}_f/A_f)$ represents the surplus time available to workers after they have worked
long enough to afford their subsistence amount of food. The expenditure shares $\Omega$ determine how much of that surplus time is allocated to each sector. For food production, we must also add back in the labor employed producing the subsistence requirement.

Given our solutions for labor allocations, we can also describe the level of urbanization in the economy. We assume that the non-tradable and tradable sectors operate within urban areas due to economies of scale and agglomeration effects. Hence the urbanization rate is

$$u = \left(1 - \frac{\tau_f}{A_f}\right) (\Omega_n + \Omega_d).$$ (22)

Our model contains within it the typical model of agricultural-led structural change and urbanization. An increase in $A_f$ will increase the surplus time available to workers, and this will raise the fraction of workers employed in the resource, non-tradable, and tradable sectors, while reducing the fraction in food. On net, as $\Omega_n$ and $\Omega_d$ are rising, urbanization will increase.

An industry-led process of structural change and urbanization can also take place. If $A_d$ increases this has several influences. First, the fraction $\Omega_f$ falls as individual substitute away from food consumption. This raises the fraction of time spent on all of the non-food goods. There will be some reallocation away from tradables towards resources and non-tradables due to the complementarity of these goods, but on net $\Omega_d$ will increase. Therefore urbanization will rise overall with industry-led productivity growth.

Note, though, that it is quite possible for resources to lead the urbanization and structural change process. An increase in $A_r$ will have similar effects to an increase in $A_d$. Labor will be drawn away from food production and into non-food production. In autarky, greater natural resource wealth can be conducive to structural change into tradable production. As we will see below, the effects of greater resource wealth change once there is the possibility of free trade.

4. OPEN ECONOMIES

Starting from autarky situation just described, what happens if we allow the economy to trade? When the economy can trade, the optimal allocations to different sectors, and hence the level of urbanization, will change according to their comparative advantage. We presume that both resources and tradables can be traded at fixed world prices, $p_r^*$ and $p_d^*$.

4.1 Comparative Advantage in Tradables

We first consider an economy where the autarky price is such that $p_r/p_d > p_r^*/p_d^*$, so the economy will export tradables and import resources. We conceive of this situation as representing the typical model of export-led development, in which a country uses a comparative advantage in producing tradables to urbanize and increase living standards. Effectively, trade allows this economy to “produce” resources in urban areas by employing tradable sector workers.

The index of non-food productivity in this economy is thus altered from that in the autarky case, and we can write the productivity index in the tradable economy as

$$A_{C}^T = \left[\frac{p_r A_d}{p_d} \theta_r + A_n^{\epsilon-1} \theta_n + A_d^{\epsilon-1} \theta_d \right]^{1/(\epsilon-1)}.$$ (23)

The productivity of this economy in turning labor into resource goods now is dependent on their tradable-sector productivity $A_d$ and the relative price of tradable and resource goods. Note that because $p_d^* A_d/p_r^* > A_r$ in this economy, trade has increased productivity and hence the real wage.
Free mobility between the three sectors that operate ensures that
\[ w = p_f A_f = p_n A_n = p_d^* A_d. \] (24)

The determination of the allocation of labor proceeds as before, only now note that the fraction of labor in the resource sector is zero, and the tradable sector must employ enough labor to supply tradable goods that can be traded for the resources demanded.

\[ \frac{L_d}{L} = \left( 1 - \frac{\bar{c}_f}{A_f} \right) \left( \Omega_d + \Omega_r \right), \] (25)

where \( \Omega_d \) and \( \Omega_r \) are defined similarly to before, only they now have \( A_C^D \) in their denominators. The expenditure fraction for resources no longer depends on domestic \( A_r \), and is instead

\[ \Omega_r = \left( \frac{p_d^* A_d}{p_r^*} \right)^{\epsilon-1} \theta_r / A_C^D. \] (26)

The fraction employed in non-tradables and food are determined as in autarky, simply noting that \( A_C^D \) replaces \( A_C \) in the denominators of \( \Omega_f \) and \( \Omega_n \).

There are two aspects of trade that lead to increased urbanization in the economy with a comparative advantage in tradable goods. First, trade has increased the productivity of the economy and this will lead to substitution away from food and into non-food goods, which includes tradables and non-tradables. Secondly, labor that would have otherwise been employed in producing resources is reallocated by trade into the tradable sector.

In the economy that specializes in tradable goods, increased urbanization is associated with structural change in the sense that workers are moving out of agriculture and into the tradable sector.

### 4.2 Comparative Advantage in Resources

We can constrain the situation of a comparative advantage in tradables to an alternative, a comparable advantage in resources. Assume that \( p_r/p_d < p_r^*/p_d^* \), so that an economy has relatively cheap resources compared to the world. It thus exports resources and imports tradable goods.

For this economy, the index of non-food productivity is now

\[ A_C^R = \left[ A_r^{\epsilon-1} \theta_r + A_n^{\epsilon-1} \theta_n + \left( \frac{p_r^* A_r}{p_d^*} \right) \theta_d \right]^{1/(\epsilon-1)}. \] (27)

The economy is able to convert labor into tradable goods at a higher rate by employing it to produce resources that can be traded for tradable goods. The productivity level \( A_C^R \) is higher than in autarky, and so the real wage has increased through trade.

The three sectors that operate employ all the labor of the economy, and again we assume free mobility between sectors so that we have

\[ w = p_f A_f = p_r A_r = p_n A_n. \] (28)

The labor allocations can be solved for as before. The resource sector must meet domestic demand as well as provide exports to acquire tradable goods, so we have

\[ \frac{L_r}{L} = \left( 1 - \frac{\bar{c}_f}{A_f} \right) \left( \Omega_r + \Omega_d \right), \] (29)

where \( \Omega_r \) and \( \Omega_d \) are defined as before, except with \( A_C^R \) in their denominators. \( \Omega_d \) is not dependent on
domestic tradable productivity, but is now

\[ \Omega_d = \left( \frac{p^* A_r}{p^*_d} \right)^{\epsilon - 1} \theta_d / A^R_C. \]  

While labor is moved out of the tradable sector and into the resource sector in this economy due to trade, there is also a tendency for higher urbanization due to the income effect of trade. Productivity in the non-food sectors has risen, and this induces a substitution away from the food sector and into non-tradables and resources. The non-tradable sector thus increases in size, which acts to raise urbanization.

In a situation where the tradable sector was very small to begin with, as was the case in Africa in the immediate post-war period, the net effect of opening to trade can be positive for urbanization. In this sense we have resource-led urbanization, but that is not associated with a structural shift towards tradable production. The urban population consists entirely of non-tradable workers.

Additionally, consider the consequences of an increase in \( A_r \), representing increased productivity in resource productivity or alternatively the discovery of new sources of resources for trade. As \( A_r \) rises, so does the productivity index \( A^R_C \). This causes a substitution away from food production and into the non-food sectors. While there also is a reallocation within the non-food sectors of labor from resources into non-tradables, it can be shown that the following holds

\[ \frac{\partial L_n}{L} / \partial A_r > 0 \quad \text{and} \quad \frac{\partial L_r}{L} / \partial A_r > 0. \]  

The result is that when \( A_R \) rises the urbanization rate increases, as shown in Figure 6. In addition, the fraction of labor employed in resources (which is also the fraction of GDP in resources in our model) also rises, and so we have a positive relationship between the urbanization rate and the size of the resource sector, as in Figure 6.

5. DYNAMICS OF STRUCTURAL CHANGE WITH TRADE

To this point we have established the effect of a one-time opening to trade on the structure of two different types of economies. Here, we extend the analysis to consider the paths of these economies over time.

To proceed, we presume that there is positive productivity growth in both the tradable and resource sectors. One can think of a learning-by-doing setting in which production of the good leads to greater skill in producing the good. For resources, however, the productivity growth is slower as there is the additional drag of declining stocks of resources. The result is that productivity growth is positive, but lower, in resources than in the tradable sector.

Positive growth in \( A_d \) in the tradable economy will lead to urbanization and a structural shift out of food production and non-tradables and towards tradable production. Figure 7 shows how urbanization and the shares in the two urban sectors change over time with productivity growth in \( A_d \). As can be seen, urbanization consists of an increasing share of tradable workers.

The resource economy experiences growth in \( A_r \), which also promotes urbanization. However, as this economy only has non-tradable production in urban areas, urbanization consists entirely of an increase in the that sector. As mentioned previously, the rise of “consumption” cities, as opposed to the production cities in the tradable-specializing economy.

Finally, we can compare the effect on living standards between the two countries. To compare the real wage between the two types of economies, we need a common price level, which we denote \( P^* \), representing some international price index. The real wages of the tradable economy relative to the resource economy can be written as follows

\[ \frac{w^R / P^*}{w^D / P^*} = \frac{p^* A^D_d}{p^*_d A^D_r}. \]  

14
and depends on the relative price of tradables to resources as well as the levels of technology in those sectors in the two economies. We presume that the relative price of tradables stays constant, as we are not modelling the equilibrium determination of that price, and it likely depends heavily on the characteristics of developed countries. The comparison of relative wages depends entirely on the path of productivity in the exporting sectors of the two economies, and as noted earlier we presume that learning-by-doing is more pronounced in the tradable economy. Hence the real wage in the tradable economy is going to advance faster than in the resource economy, despite both beginning to urbanize. Figure 8 shows the increase in relative wages in the tradable economy relative to the resource exporter.

6. CONCLUSION

This paper develops a new model of structural transformation in open economies that can help account for the divergent patterns of urbanization and development observed in Asia and Africa. Asia has experienced urbanization with structural transformation and growth in incomes; in our model, this reflects their success in exporting tradable manufactured goods and services to the rest of the world. Since these tradables are produced in cities, comparative advantage and specialization drive urban growth through the rise of production cities.

In contrast, we argue that Africa has experienced urbanization without structural transformation. African countries export natural resources to the rest of the world. Resource revenues are spent in the cities, whether by private recipients of resource rents or by governments. The income effect drives urban growth through the rise of consumption cities.

If the production of tradable manufactured goods and services requires more skilled – and higher paid – workers, then Asia's pattern of urbanization will be associated with development, while Africa's pattern will lead to urbanization with little income growth.

Our model makes three contributions to our understanding of the relationship between structural change, urbanization and economic development. The first contribution is to establish that Africa has followed a different pattern of urbanization from elsewhere. We argue that this is because Africa did not urbanize following a Green Revolution or an Industrial Revolution, but as a result of natural resource exports. The second contribution is to question the supposedly positive relationship between urbanization and economic development. Although resource-rich countries are much more urbanized than resource-poor countries within Africa, resource-rich countries are over-urbanized for their level of economic development. We explain this result by the fact that African cities developed without producing tradable goods and without human capital accumulation. The third contribution is to highlight a new mechanism for the resource curse. Resource-rich countries urbanize but remain poor.

This paper leaves several open questions. The first is why African countries have been unable to acquire or develop a comparative advantage in sectors other than those based on resource extraction. We acknowledge the possibility that institutions and colonial history, as well as resource endowments, may have driven this initial specialization. But we do not attempt to model this directly. A second question that we leave unanswered is why consumption cities do not evolve into production cities over time. Our model does not allow for this, and we view it as an important question that warrants further study.
REFERENCES


FIGURES

Figure 1: Income and Urbanization, Asia versus Africa, 1950-2010.

Sources: Maddison 2008, UN 2011, WDI 2011, USGS 2011, authors’ calculations. Per capita GDP is reported in 1990 Geary-Khamis dollars (constant, PPP). The urbanization rate is reported using national urban definitions. The data includes 22 Asian countries (Eastern, South-Eastern and South Asia) and 41 Sub-Saharan African countries.

Figure 2: Urbanization and Composition of GDP, Asia versus Africa, 2000.

Sources: WDI 2011, USGS 2011, FAO 2011, UN 2011, authors’ calculations. The urbanization rate is reported using national urban definitions. GDP is decomposed into manufacturing and services, resource exports (fuel, mining and cash crop exports) and agriculture (for domestic consumption). The GDP share of agriculture is not reported. The data includes 22 Asian countries (Eastern, South-Eastern and South Asia) and 41 Sub-Saharan African countries. We display in red the quadratic prediction plots.
Figure 3: Income, Resource Exports and Urbanization in Africa, 1950-2010.

Sources: Maddison 2008, UN 2011, WDI 2011, USGS 2011, authors’ calculations. Per capita GDP is reported in 1990 Geary-Khamis dollars (constant, PPP). The urbanization rate is reported using national urban definitions. The data includes 22 Asian countries (Eastern, South-Eastern and South Asia) and 41 Sub-Saharan African countries. African countries are separated into four groups, depending on the average GDP share (%) of natural resource exports in 1960-2000.
Figure 4: Urbanization for Eight Resource-Rich African Countries, 1950-2010.

Sources: UN 2009, authors’ calculations. The urbanization rate is reported using national urban definitions. The eight resource-rich African countries are Gabon (oil), Angola (oil), Nigeria (oil), Liberia (iron ore, gold and diamonds), Ghana (cocoa and gold), Ivory Coast (cocoa and coffee), Zambia (copper) and Botswana (diamonds).
Figure 5: Urbanization for Eight Resource-Poor African Countries, 1950-2010.

Sources: UN 2009, authors’ calculations. The urbanization rate is reported using national urban definitions. The eight resource-poor African countries are Burundi, Chad, Ethiopia, Lesotho, Rwanda, Swaziland, Uganda and Niger.
Figure 6: Comparative Advantage in Resources

Figure 7: Comparative Advantage in Tradables

Figure 8: Comparison of Real Wages over Time