Schooling Choice during Structural Transformation

by

Paul Schreck

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved April 2011 by the
Graduate Supervisory Committee:
Berthold Herrendorf, Chair
David Lagakos
Todd Schoellman

ARIZONA STATE UNIVERSITY
May 2011
ABSTRACT

This dissertation consists of two essays. The first measures the degree to which schooling accounts for differences in industry value added per worker. Using a sample of 107 economies and seven industries, the paper considers the patterns in the education levels of various industries and their relative value added per worker. Agriculture has notably less schooling and is less productive than other sectors, while a group of services including financial services, education and health care has higher rates of schooling and higher value added per worker. The essay finds that in the case of these specific industries education is important in explaining sector differences, and the role of education all other industries are less defined.

The second essay provides theory to investigate the relationship between agriculture and schooling. During structural transformation, workers shift from the agriculture sector with relatively low schooling to other sectors which have more schooling. This essay explores to what extent changes in the costs of acquiring schooling drive structural transformation using a multi-sector growth model which includes a schooling choice. The model is disciplined using cross country data on sector of employment and schooling constructed from the IPUM International census collection. Counterfactual exercises are used to determine how much structural transformation is accounted for by changes in the cost of acquiring schooling. These changes account for small shares of structural transformation in all economies with a median near zero.
DEDICATION

To Lisa for her support.
ACKNOWLEDGEMENTS

I am deeply indebted to Berthold Herrendorf, David Lagakos and Todd Schoellman for their advice and encouragement. I would also like to thank Kevin Donovan, Daniel Lawver, Richard Rogerson, Loris Rubini, Johanna Wallenius and the seminar participants at Arizona State University, for helpful discussions and comments. All remaining errors are my own.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>TABLE OF CONTENTS</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIST OF TABLES</td>
<td>vi</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>vii</td>
</tr>
<tr>
<td>CHAPTER</td>
<td>1</td>
</tr>
<tr>
<td>1 EDUCATIONAL ATTAINMENT AND INDUSTRY PRODUCTIVITY</td>
<td>1</td>
</tr>
<tr>
<td>1.1 Introduction</td>
<td>1</td>
</tr>
<tr>
<td>1.2 Environment</td>
<td>3</td>
</tr>
<tr>
<td>1.3 Data</td>
<td>5</td>
</tr>
<tr>
<td>1.4 Findings</td>
<td>9</td>
</tr>
<tr>
<td>1.5 Conclusion</td>
<td>20</td>
</tr>
<tr>
<td>2 SCHOOLING CHOICE DURING STRUCTURAL TRANSFORMATION</td>
<td>21</td>
</tr>
<tr>
<td>2.1 Introduction</td>
<td>21</td>
</tr>
<tr>
<td>2.2 Motivation</td>
<td>26</td>
</tr>
<tr>
<td>2.3 Model</td>
<td>29</td>
</tr>
<tr>
<td>3.1 Equilibrium</td>
<td>33</td>
</tr>
<tr>
<td>3.2 Household problem</td>
<td>34</td>
</tr>
<tr>
<td>3.3 Agriculture Firms problem</td>
<td>34</td>
</tr>
<tr>
<td>3.4 Non-agriculture Firms problem</td>
<td>34</td>
</tr>
<tr>
<td>3.5 Education Firms problem</td>
<td>35</td>
</tr>
<tr>
<td>Chapter</td>
<td>Page</td>
</tr>
<tr>
<td>---------</td>
<td>------</td>
</tr>
<tr>
<td>Market Clearing conditions</td>
<td>35</td>
</tr>
<tr>
<td>2.4 Data and Calibration</td>
<td>36</td>
</tr>
<tr>
<td>Data</td>
<td>36</td>
</tr>
<tr>
<td>Calibration</td>
<td>37</td>
</tr>
<tr>
<td>Endowments</td>
<td>37</td>
</tr>
<tr>
<td>Preferences</td>
<td>38</td>
</tr>
<tr>
<td>Technology</td>
<td>38</td>
</tr>
<tr>
<td>Economy Calibration</td>
<td>42</td>
</tr>
<tr>
<td>Calibration Example: Brazil</td>
<td>44</td>
</tr>
<tr>
<td>2.5 Findings</td>
<td>45</td>
</tr>
<tr>
<td>Counterfactual Exercise 1</td>
<td>46</td>
</tr>
<tr>
<td>Counterfactual Exercise 2</td>
<td>46</td>
</tr>
<tr>
<td>Analysis</td>
<td>47</td>
</tr>
<tr>
<td>Comparison to Caselli and Coleman</td>
<td>50</td>
</tr>
<tr>
<td>2.6 Conclusion</td>
<td>53</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>54</td>
</tr>
</tbody>
</table>
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Industry Schooling</td>
</tr>
<tr>
<td>1.2</td>
<td>Industry Value Added per Worker</td>
</tr>
<tr>
<td>1.3</td>
<td>Baseline $m_{j,ag}^1$</td>
</tr>
<tr>
<td>1.4</td>
<td>Baseline $m_{j,ag}^2$</td>
</tr>
<tr>
<td>1.5</td>
<td>Alternative $m_{j,ag}^2$</td>
</tr>
<tr>
<td>1.6</td>
<td>Baseline $m_{j,man}^1$</td>
</tr>
<tr>
<td>1.7</td>
<td>Baseline $m_{j,man}^2$</td>
</tr>
<tr>
<td>1.8</td>
<td>Alternative $m_{j,man}^2$</td>
</tr>
<tr>
<td>1.9</td>
<td>Counterfactual aggregate value added per worker relative to data</td>
</tr>
<tr>
<td>2.1</td>
<td>Schooling Model for Brazil</td>
</tr>
<tr>
<td>2.2</td>
<td>Findings with different elasticities</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1 $m_{ag,man}^1$ histogram: number of economies with each value of $m_{ag,man}^1$</td>
<td>14</td>
</tr>
<tr>
<td>1.2 $m_{ag,man}^2$ histogram: number of economies with each value of $m_{ag,man}^2$</td>
<td>15</td>
</tr>
<tr>
<td>1.3 $m_{OA,man}^2$ histogram: number of economies with each value of $m_{OA,man}^2$</td>
<td>17</td>
</tr>
<tr>
<td>1.4 African economies: A comparison of industry efficiency units per worker and value added per worker normalized by the agriculture industry</td>
<td>18</td>
</tr>
<tr>
<td>2.1 Schooling change and structural transformation for 11 economies</td>
<td>27</td>
</tr>
<tr>
<td>2.2 Share of schooled workers in agriculture and non-agriculture sectors</td>
<td>28</td>
</tr>
<tr>
<td>2.3 Share of schooled workers in education and non-agriculture sectors</td>
<td>29</td>
</tr>
<tr>
<td>2.4 Non-agriculture and agriculture log schooling shares for sample countries</td>
<td>41</td>
</tr>
<tr>
<td>2.5 Non-agriculture and education log schooling shares for sample countries</td>
<td>42</td>
</tr>
<tr>
<td>2.6 Labor types and Sectors in Brazil, Model and Data compared</td>
<td>44</td>
</tr>
<tr>
<td>2.7 Findings from Counterfactual Exercise 1: Changes explained with no change in schooling parameters</td>
<td>47</td>
</tr>
<tr>
<td>2.8 Findings for Counterfacutal Exercise 2: Change of initial economies from a change of +50% TFP education, and -50% $\psi$</td>
<td>48</td>
</tr>
<tr>
<td>2.9 Changes in relative productivity of agriculture and structural transformation</td>
<td>49</td>
</tr>
<tr>
<td>2.10 Subsistence explaining Structural Transformation from Schooling, counterfactual 2</td>
<td>50</td>
</tr>
<tr>
<td>2.11 Findings for Counterfacutal Exercise 2 for different technology parameterizations: Change of initial economies from a change of +50% TFP education and -50% $\psi$</td>
<td>52</td>
</tr>
</tbody>
</table>
Large differences in income between countries continues to be a major puzzle for economists. There are vast differences in output per worker between the lowest income economies and economic leaders. The average American worker has 30 times the real output of a Rwandan worker. One potential candidate for explaining this difference is differences in skills acquired through education. It is reasonable to think that basic skills such as math, reading and science would increase a worker’s productivity in a wide variety of tasks. The data shows large differences in educational attainment in rich and poor economies. 80% of Rwandan workers have little or no primary school education, while 65% of American workers have a high school diploma in the year 2000 and 85% have completed primary school. Low income economies also have very different industry structures than high income economies. Specifically, the share of workers in agriculture is much higher in low income economies. This is important because agriculture value added per worker in low income economies is worse than value added per worker in other industries when compared to high income economies. Low income economies seem to be weighed down by large unproductive agriculture industries.

I investigate education, industry productivity and their interaction. Does looking at education for individual industries provide greater insight on why some industries are more productive? Many economists including Bils and Klenow (AER 2003), Hall and Jones (QJE 1999) and Klenow and Rodriguez-Clare(1997) have measured the contribution of education in explaining cross country differences in productivity. I will apply similar methodology for new data on educational attainment for industry’s participants. This will show how much education accounts for differences in industry productivity within 107 economies. This provides some
insight on the relative importance of education policy for targeting economic productivity. This analysis will provide information about the effects of education, industry productivity and industry composition on aggregate productivity.

The contribution of this paper begins with the unique data aggregation: schooling information by industries in the economy for a number of low income economies and times series data for growing economies in combination with value added for industry from national accounts. This data opens a gateway to understanding the interactions of structural transformation and education in the post war era. One striking fact is the consistent ordering of major industries by education: transportation workers are more educated than construction workers which, in turn, are more educated than agriculture workers. After considering data observations, I calculate contributions of schooling in determining differences in industry productivity. This is done by comparing value added per worker in various industries and education adjusted efficiency per worker for those industries. Education explains roughly 15% of industry productivity differences between agriculture and other industries.

My analysis will focus on seven major industries: agriculture, construction, manufacturing, wholesale and retail, transportation and communication, mining and utilities, and other services. While further disaggregation is possible from education data, these industries allow for comparison with a wide variety of economies. There are 44 countries and 107 total censuses which include high income countries, several Latin American countries, China, India, Southeast Asian countries and six African nations. Available censuses between 1970 and 2005 are used in this paper. This is useful because many economies have undergone significant structural changes over this period. It is not possible to exaggerate thanks to the Integrated Public Use Microdata Series (IPUMS) International for their work in providing consistent microdata from which the industry aggregates in
this paper are constructed.

The paper will present data and counterfactuals for Africa, Latin America and the economies with the largest agriculture employment. Education does not explain much of the large disparities between agriculture and other industries, however, we find that there can be significant gains to aggregate productivity by increasing education closer to that of the most educated industries. The intent of this paper is determine in which industries education is important in explaining productivity differences exist between industries and which industries leave the largest unexplained value added per worker differences. I conclude that the most interesting industries in this respect are agriculture and "other services".

I begin in section 2 presenting the environment, specifically the production technology used for measurement. Section 3 documents the data that was used and discuss the validity of the measures. Discussion of the details of measurement and findings are presented in section 4. I conclude with a discussion of implications of these findings.

1.2 Environment

Productivity will be defined with a simple measure of value added per worker. I will abstract from industry differences in capital and natural resource. Consider an economy with a production technology for each industry which is linear in labor input and a standard Mincer return to schooling:

\[ f_j(\{N_{ij}\}_{i=1}^{I}) = A_j \sum_{i=1}^{I} e^{\beta_i s_i} N_{ij} \]

In this production function, \( j \) indexes industry and \( i \) indexes the level of educational attainment. \( N_{ij} \) denotes the number of workers in an industry with a level of educational attainment \( i \). \( N_j \) denotes the total number of workers in a given industry:

\[ N_j = \sum_{i=1}^{I} N_{ij} \]
The elements in the Mincer efficiency formulation are the Mincer coefficient $\beta_i$ which is the return to schooling and $s_i$, the number of years of schooling corresponding to the level of educational attainment $i$. For example, $N_{a2}$ would denote the number of agriculture workers with at most primary school completed, so $s_2 = 6$.

I will call the composite of Mincer efficiency and worker’s "efficiency unit" $L_j$, different from "workers" $N$:

\[
L_{ij} = e^{\beta_i s_i} N_{ij}
\]

\[
L_j = \sum_{i=1}^{I} L_{ij}
\]

Now I will differentiate between "value added per worker" or $\frac{Y_j}{N_j}$ and "value added per efficiency unit" $\frac{Y_j}{L_j}$ which accounts for productivity from education. When comparing value added per worker across industries, the use of value added per efficiency unit will correct for differences in educational attainment. The measure of value added per worker contains differences due to education and other productivity differences:

\[
\frac{Y_j}{N_j} = A_j \sum_{i=1}^{I} e^{\beta_i s_i} \frac{N_{ij}}{N_j}
\]

Compared to value added per efficiency unit:

\[
\frac{Y_j}{L_j} = A_j
\]

I can also compose the industry value added per worker and construct the total value added per worker. This will be useful for bridging the discussion of industry productivity and aggregate productivity which include compositional differences. The three major components will be the labor composition of industries, output per efficiency and worker efficiency units:

\[
\frac{Y}{N} = \sum_{j=1}^{J} A_j \frac{L_j N_j}{N_j N}
\]
This straightforward environment will allow basic patterns to emerge from the data, assisting to understand how industry productivity, education and industry shares interact in creating the huge aggregate productivity differences we observe in the data.

1.3 Data

I use several sources for data, the most important being the Integrated Public Use Microdata Series (IPUMS) from the University of Minnesota which includes census data for several countries. This is my source for educational attainment and industry participation data. I employ National Accounts reports from the National Accounts Main Aggregate Database to find the industry value added. I use the research of Psacharopoulos and Patrinos (1994, 2004) and Schoellman (2009) for measures of returns to schooling. In this section I will discuss each of these three data sources, providing an overview of what is available and what I employ in my measurement.

IPUMS

The Integrated Public Use Microdata Series has census data from a large number of economies from as early as 1970. I will use data from 44 countries including 107 censuses where data is available on schooling, industry, employment and corresponds to dates where national accounts are also available. A full list of countries, census dates and major aggregates are available in the appendix. They include low income countries from Latin America, Europe, Africa and East Asia, as well as middle and high income countries. The sample has a large representation of Latin American countries, 11 countries all with multiple censuses. Within this region there is structural diversity from agrarian Brazil of 1970 to emerging Chile of 2002. Much of the world’s population is included in India and China. I will weight all results equally regardless of population because those economies would dominate any weighting scheme. While OECD countries are perhaps
underrepresented, I have 8 countries and 27 census in the sample.

While most countries include a relatively recent census (1990-2002) several countries provide a series of censuses from the post war era. Because countries like Brazil, France and the United States change dramatically in terms of education and relative industry size, I will consider each census as an individual economy. Censuses are infrequent, typically 10 years apart. The sample sizes are large, typically 2-10% of households with the exception being China (1983) which reports .5% of households and India which has .09% in employment surveys. IMUPS has many variables including industry, employment status, educational attainment, occupation, homeownership, sex, age, marriage status, children, region and other typical census questions. Wage and hours information would be very useful, however it is not available for almost all economies. I will restrict my observations to employed workers who report educational attainment and industry.

All information in the dataset has been standardized for ease of comparison. While this is occasionally problematic and requires a closer examination of underlying data, it simplifies the aggregation of the millions of observations. I will restrict my interest in industries to seven industries: Agriculture, hunting, forestry, fishing (ISIC A-B), Mining, utilities (ISIC C,E), Manufacturing (ISIC D), Construction (ISIC F), Wholesale, retail trade, restaurants and hotels (ISIC G-H), Transport, storage and communication (ISIC I), Other Activities (ISIC J-P) . Other Activities (ISIC J-P) include education, financial services, insurance, real estate, business services, health, social work, public administration, defense, private household services and other services. The "other activities" category includes many of the industries associated with higher education: education, medical, legal and financial services. These categories were taken from the United Nations Statistics National Accounts data. While further disaggregation may be feasible, these seven common industries are important for development
The IPUMS provides several representations of educational attainment including years of schooling for various economies and highest level completed. Because highest level completed is the most available measure, I used that measure of attainment for consistency across samples. I have four categories of educational attainment: no primary completed, primary school completed, secondary school completed and University completed; representing the highest achieved level of schooling for the individual. These will correspond to the standard \{0,6,12,16\} years of schooling. While the exact level may vary country-to-country or year-to-year, these are consistent approximations for years associated with various levels of education.

Value Added

As stated above, the value added data for industries comes from the National Accounts Main Aggregate Database made available by United Nations Statistical Division. These are the National Accounts reported to the United Nations Statistical Division and are verified with data from countries statistics agencies, central banks, the IMF and OECD. The relative prices are local. Because of the wide variety of problems of industry pricing across economies, I will focus on comparisons of industries with a single economy. There are a few national accounts which aggregate slightly differently than listed above. In these cases, labor aggregations were adjusted to be consistent. The database has an extensive list of countries extending with timeseries back to 1970 for most countries. While the quality of the data report by low income economy statistical offices is always suspect, the existence of a publically available national census in the same year provides some control for competent government. Almost all economies after 1970 for which a census existed had value added data available. A notable exception is the Palestinian economy. Special thanks to the economists at the United Nations
Statistical Division for making this data available.

There are composition concerns with some of these industries. Manufacturing includes a greater number of handcrafts in low income economies and microchips in high income economies. Similarly, the composition of Other Activities can be very different for high income economies and low income economies. Defining goods and industries is always difficult, but the nature of the aggregation makes some sectors more homogenous like Construction or Agriculture while Manufacturing and Other Activities are less homogenous. This is a data limitation which can only be addressed by a greater disaggregation of industries which is currently not available for a large number of economies. Since these same problems exist to a greater degree with economy aggregates or standard Agriculture, Manufacturing, Services industries, the aggregation provided in this paper is a big step forward in separating composition effects from industry specific effects.

*Returns to Schooling*

Psacharopoulos and Patrinos (1994, 2004) and Bils and Klenow (2000) provide numerous estimates from microdata studies for wage returns to schooling. The papers all use a Minser formulation to estimate the coefficient of years of schooling on log wages. These estimates are for a wide variety of economies and vary from 4% to 20% returns to a year of school. Unfortunately, the methodology and results of these studies are not consistent. One finds significant differences in similar studies conducted only a year apart. For this reason, there is reason to be skeptical of any one estimate. Only a small subset of economies in my data set have a returns to schooling measurement within 5 years of the census date, therefore these studies will not be directly applicable. Surprising is the finding that the return to schooling is not highly correlated to income or other macro factors. In testing correlations, a regional dummy variable was most significant.
Because the data does not present a clear solution to finding country specific returns to schooling coefficients, I will use a simple baseline and a more conservative alternative to check robustness. The baseline will use the value 10% which is very close to the mean, median and value used for the United States. The alternative case is a higher returns to schooling consistent with Schoellman (2009) which can account for quality differences. This 20% return to schooling acts as an upper bound for my findings. Despite an extensive literature on Minser coefficients, there is not consensus of which returns to education values are best in a macroeconomic setting. Since considerations extend from quality differences to selections to general equilibrium effects, this paper will take a simple approach within the tradition of the literature and an aggressive alternative. I have checked the Minser regressions for several of the censuses, which confirm the findings of the literature: large differences between economies but roughly consistent with 10% returns per year of schooling.

1.4 Findings

I will represent the findings in three parts. First, I will discuss the ordering of industries found in the data, describing “most educated industries” and “most value added industries”. This provides some major patterns emerging from the data and a sense of the findings. Then I will provide two measures of the contribution of educational attainment to relative productivity of industries. Agriculture and manufacturing are chosen as base industries and other industries are presented relative to those industries. The final section will be a counterfactual exercise comparing the contribution of factors in explaining aggregate total value added. What would an economy’s productivity look like if education or industry composition looked like the most educated industry?
Industry Education and Productivity

Which industries are most educated? Which industries are most productive? While one may have some beliefs about the education of those working in an industry, these patterns may vary from economy to economy. I used three measures of education to create this ordering: \( \frac{L_j}{N_j} \) using the two versions of return to education and share of workers with at least secondary education. The orderings are slightly different for some economies, but the major findings are consistent for all three approaches.

There are three strong claims which I can make from this data. First, agriculture workers have the least education by a large margin. The closest industry has, on average, 20% more \( \frac{L_j}{N_j} \). The second finding is that construction is the second least educated industry. Construction is either least educated or second least educated in 86% of economies in the baseline case. The third finding is that Other Activities which includes many of the high skilled industries has significantly higher education than all other industries. Other economies is the most educated economy in 93% of economies in the baseline case. Average \( \frac{L_j}{N_j} \) is 20% higher than the next highest industry in the baseline and 10% higher in the alternative. Table 1.1 provides the complete ordering.

Table 1.1: Industry Schooling

<table>
<thead>
<tr>
<th>Industry</th>
<th>Median share with secondary</th>
<th>Mean Baseline ( \frac{L_j}{N_j} )</th>
<th>Median Baseline ( \frac{L_j}{N_j} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agriculture</td>
<td>5%</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Construction</td>
<td>15%</td>
<td>1.21</td>
<td>1.2</td>
</tr>
<tr>
<td>Manufacturing</td>
<td>21%</td>
<td>1.33</td>
<td>1.33</td>
</tr>
<tr>
<td>Wholesale, Retail...</td>
<td>26%</td>
<td>1.36</td>
<td>1.36</td>
</tr>
<tr>
<td>Transportation,...</td>
<td>28%</td>
<td>1.39</td>
<td>1.4</td>
</tr>
<tr>
<td>Mining, Utilities</td>
<td>31%</td>
<td>1.45</td>
<td>1.43</td>
</tr>
<tr>
<td>Other Activities</td>
<td>50%</td>
<td>1.72</td>
<td>1.74</td>
</tr>
</tbody>
</table>

The ordering of Manufacturing, Wholesale, retail trade, restaurants and
hotels, Transport, storage and communication, Mining and utilities is reasonably robust. $L_i / N_i$ for Wholesale, retail trade, restaurants and hotels exceeds Manufacturing in 66% of the economies. $L_i / N_i$ for Transport, storage and communication exceeds Wholesale, retail trade, restaurants and hotels in 71% of economies. $L_i / N_i$ for Mining and utilities exceeds Transport, storage and communication in 65% of economies for the baseline and 56% of economies for the alternative.

The industry with the highest value added per worker is Mining and Utilities and the lowest highest value added per worker is Agriculture. Mining and Utilities are industries with high capital and resources which are measured as higher value added per worker. If the existing value of the land with oil or ore were better calculated as an input in mining, the value added per worker is possibly lower. It is readily obvious that schooling is not a large factor in explaining Mining and Utilities value added per worker. Agriculture is the least educated industry and the least productive industry. The differences between Agriculture and other industries is much greater in value added per worker than education measures. There are some exceptions including Argentina, Iraq, Israel and the United States. Argentina has a highly developed agriculture sector and is often an outlier in cross country studies of agriculture. Iraq and Israel are interesting because of its unique climate and ecology, which may require more capital intensive agriculture. The United States in 2005 has greater value added per worker in Agriculture than Construction and about the same as Wholesale, retail trade, restaurants and hotels. The remaining sectors cannot be clearly ordered, instead there are two groups. The lower value added per worker group is Manufacturing, Wholesale, retail trade, restaurants and hotels, and Construction, which are the less educated of the remaining sectors. The higher value added per worker group is Transportation, communication and Other activities. Transportation and communication are
potentially higher capital industries, so value added per worker should be higher.

Table 1.2: Industry Value Added per Worker

<table>
<thead>
<tr>
<th>Industry</th>
<th>Mean $\frac{Y_j}{N_j}$</th>
<th>Median $\frac{Y_j}{N_j}$</th>
<th>1st Quartile $\frac{Y_j}{N_j}$</th>
<th>3rd Quartile $\frac{Y_j}{N_j}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agriculture</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Manufacturing</td>
<td>3.93</td>
<td>2.72</td>
<td>1.97</td>
<td>4.99</td>
</tr>
<tr>
<td>Wholesale, Retail</td>
<td>3.56</td>
<td>2.88</td>
<td>1.75</td>
<td>4.44</td>
</tr>
<tr>
<td>Construction</td>
<td>4.19</td>
<td>2.94</td>
<td>1.66</td>
<td>4.5</td>
</tr>
<tr>
<td>Transportation,...</td>
<td>4.56</td>
<td>3.54</td>
<td>2.13</td>
<td>5.83</td>
</tr>
<tr>
<td>Other Activities</td>
<td>4.46</td>
<td>3.83</td>
<td>2.37</td>
<td>5.94</td>
</tr>
<tr>
<td>Mining, Utilities</td>
<td>20.15</td>
<td>9.1</td>
<td>5.41</td>
<td>23.65</td>
</tr>
</tbody>
</table>

Contribution of Education

Here I will consider the share of value added account for by education. First I consider the value added per worker of a industry relative to the index industry, either Agriculture or Manufacturing $\frac{Y_j}{N_j}$. From $\frac{L_j}{N_j}$ we can create a model value of $\frac{Y_j}{N_j}$ assuming that the output per efficiency units $A$ is the same in both sectors:

$$\frac{Y_{j,\text{model}}}{N_{j,\text{model}}} = \frac{A}{N_j} \frac{L_j}{N_j} = \frac{Y_j}{N_j}$$

The difference between a model and data can be expressed by the ratio:

$$m_{j,k}^1 = \frac{\frac{Y_{j,\text{model}}}{N_{j,\text{model}}} - \frac{Y_{k,\text{model}}}{N_{k,\text{model}}}}{\frac{Y_j}{N_j} - \frac{Y_k}{N_k}}$$

While this measure is useful when $\frac{Y_j}{N_j} > \frac{Y_k}{N_k}$ and $\frac{L_j}{N_j} > \frac{L_k}{N_k}$, there are many cases of sector pairs for which $\frac{Y_j}{N_j} > \frac{Y_k}{N_k}$ but $\frac{L_j}{N_j} < \frac{L_k}{N_k}$. According to this metric, $m_{j,k}^1$ can be quite large if $\frac{Y_j}{N_j}$ is close to 1 even when $\frac{L_j}{N_j} < \frac{L_k}{N_k}$. Therefore, an alternative metric is necessary for our industry comparison. I propose $m_{j,k}^2$ which is based on the share of the percentage "increase" of the industry value added per worker relative to the percentage "increase" in the model value added per worker. So, for
$Y_j > Y_k$, the metric $m_{j,k}^2$ would be:

$$m_{j,k}^2 = \frac{Y_j, model \cdot Y_k, model - Y_k, model \cdot Y_j, model}{Y_j, model \cdot Y_k, model - Y_j, model \cdot Y_k, model} = 1$$

This approach is more informative when $\frac{Y_j}{N_j}$ is close to 1, but it is not standard for the development literature. To conform to established measures, I will present a modified version of $m_{j,k}^1$ where $\frac{L_j}{N_j} < \frac{L_k}{N_k}$ cases are simply set to zero. I will also provide the share of cases where $\frac{L_j}{N_j} > \frac{L_k}{N_k}$ for $\frac{Y_j}{N_j} > \frac{Y_k}{N_k}$ since these are the cases where the model predicts in the correct direction. All of these are unitless measures, but industries used for comparison is relevant. There are 21 possible combinations, and I present a subset of 11 focusing on the Agriculture and Manufacturing industries.

<table>
<thead>
<tr>
<th>Industry</th>
<th>Mean $m_{j,ag}^1$</th>
<th>$L_j &gt; L_k$ cases</th>
<th>Median</th>
<th>1st Quartile</th>
<th>3rd Quartile</th>
</tr>
</thead>
<tbody>
<tr>
<td>Construction</td>
<td>.39</td>
<td>87%</td>
<td>.34</td>
<td>.18</td>
<td>.54</td>
</tr>
<tr>
<td>Manufacturing</td>
<td>.46</td>
<td>96%</td>
<td>.41</td>
<td>.27</td>
<td>.62</td>
</tr>
<tr>
<td>Wholesale,Retail,...</td>
<td>.46</td>
<td>90%</td>
<td>.43</td>
<td>.24</td>
<td>.59</td>
</tr>
<tr>
<td>Transportation,...</td>
<td>.45</td>
<td>99%</td>
<td>.37</td>
<td>.24</td>
<td>.61</td>
</tr>
<tr>
<td>Mining, Utilities</td>
<td>.18</td>
<td>98%</td>
<td>.14</td>
<td>.06</td>
<td>.26</td>
</tr>
<tr>
<td>Other Activities</td>
<td>.49</td>
<td>97%</td>
<td>.42</td>
<td>.29</td>
<td>.66</td>
</tr>
</tbody>
</table>

Table 1.3 uses the $m_{j,k}^1$ with the adjustment of cases where the model predicts the wrong direction set to zero, comparing Agriculture to other industries using the baseline return to schooling value. As seen in the discussion of education and value added per worker in sectors, Agriculture is almost always less educated and less productive. This is visible in the $\frac{L_j}{N_j} > \frac{L_k}{N_k}$ measure which holds in almost all cases. The metric $m_{j,k}^1$ is highest for Manufacturing, Wholesale, retail, etc., Transportation and Other Activities. Schooling seems to be slightly less relevant in explaining the difference in productivity between Agriculture and...
Figure 1.1: $m_{ag,man}^1$ histogram: number of economies with each value of $m_{ag,man}^1$

Construction, and much less relevant in explaining the difference between Mining and utilities, and Agriculture. The tables include values for the first quartile, median and third quartile to provide a sense of the distribution of the values. The difference between sectors varies widely from economy to economy, but the industries seem to have similar distributions.

The histogram provides an example of the distribution of values of $m_{j,k}^1$. The distribution is quite flat between zero and .7. Table 1.4 provides analysis of the same comparison using the preferred measure $m_{j,ag}^2$. While this metric improves on the last, the mean is sensitive to outliers so one should focus on the Median, 1st Quartile and 3rd Quartile. While we see many of the broad patterns as in the $m_{j,k}^1$ case, the values are more conservative and the distributions are tighter. Other Activities also differentiates itself for having a higher share of value added per worker explained by education relative to other industries. By this measure, Construction looks much more like Mining and utilities. The corresponding frequency graph for agriculture and manufacturing, Figure 2, has a significantly
tighter distribution and a long right hand tail. The key finding is that higher manufacturing schooling levels accounts for around 13% of the higher output levels.

Table 1.4: Baseline $m^2_{j,ag}$

<table>
<thead>
<tr>
<th>Industry</th>
<th>$m^2_{j,ag}$</th>
<th>Median</th>
<th>1st Quartile</th>
<th>3rd Quartile</th>
</tr>
</thead>
<tbody>
<tr>
<td>Construction</td>
<td>-.84</td>
<td>.07</td>
<td>.02</td>
<td>.15</td>
</tr>
<tr>
<td>Manufacturing</td>
<td>.36</td>
<td>.13</td>
<td>.07</td>
<td>.26</td>
</tr>
<tr>
<td>Wholesale,Retail...</td>
<td>.83</td>
<td>.16</td>
<td>.06</td>
<td>.29</td>
</tr>
<tr>
<td>Transportation,...</td>
<td>.28</td>
<td>.15</td>
<td>.07</td>
<td>.25</td>
</tr>
<tr>
<td>Mining, Utilities</td>
<td>.07</td>
<td>.04</td>
<td>.02</td>
<td>.1</td>
</tr>
<tr>
<td>Other Activities</td>
<td>.4</td>
<td>.25</td>
<td>.14</td>
<td>.4</td>
</tr>
</tbody>
</table>

This was the conservative baseline measure of returns to schooling. Many believe this is an low estimate so I will use Schoellman’s high estimate for returns to schooling and find similar but slightly better results. Table 1.5 has these results. With this higher valuation of schooling the share of value added per worker differences explained are about twice as high. While schooling is now more important in explaining industry productivity differences, it is less than half for the...
majority economies and industries.

<table>
<thead>
<tr>
<th>Mean $m_{j,ag}^2$</th>
<th>Median</th>
<th>1st Quartile</th>
<th>3rd Quartile</th>
</tr>
</thead>
<tbody>
<tr>
<td>Construction</td>
<td>-1.54</td>
<td>.17</td>
<td>.06</td>
</tr>
<tr>
<td>Manufacturing</td>
<td>.79</td>
<td>.31</td>
<td>.18</td>
</tr>
<tr>
<td>Wholesale,Retail</td>
<td>1.71</td>
<td>.35</td>
<td>.16</td>
</tr>
<tr>
<td>Transportation,...</td>
<td>.67</td>
<td>.35</td>
<td>.18</td>
</tr>
<tr>
<td>Mining, Utilities</td>
<td>.168</td>
<td>.1</td>
<td>.04</td>
</tr>
<tr>
<td>Other Activities</td>
<td>1.11</td>
<td>.69</td>
<td>.45</td>
</tr>
</tbody>
</table>

While the comparisons with agriculture are most interesting for development questions, it is only a small segment of the industry pairs which can be compared. The second group of tables will report the metrics for Manufacturing and other industries. Here the predict direction of schooling difference and value added per worker difference is less frequent as shown in the $\frac{L_j}{N_j} > \frac{L_k}{N_k}$ statistic. As discussed earlier, Manufacturing, Construction, Wholesale, retail, etc. and to some extent, Transportation and communication were similar in terms of schooling and value added per worker. Therefore, it is not surprising to see relatively lower rates of prediction in that group. Table 1.7 shows very little prediction using the $m_{j,man}^2$ metric. Mining and utilities is almost always more productive and more educated than manufacturing, but the productivity differences are large and the education differences are relatively small, so the difference is quantitively small by all measures.

<table>
<thead>
<tr>
<th>Mean $m_{j,man}^1$</th>
<th>$\frac{L_j}{N_j} &gt; \frac{L_k}{N_k}$</th>
<th>Median</th>
<th>1st Quartile</th>
<th>3rd Quartile</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agriculture</td>
<td>.46</td>
<td>96%</td>
<td>.41</td>
<td>.27</td>
</tr>
<tr>
<td>Construction</td>
<td>.33</td>
<td>52%</td>
<td>.14</td>
<td>0</td>
</tr>
<tr>
<td>Wholesale,Retail...</td>
<td>.39</td>
<td>58%</td>
<td>.44</td>
<td>0</td>
</tr>
<tr>
<td>Transportation,...</td>
<td>.4</td>
<td>58%</td>
<td>.36</td>
<td>0</td>
</tr>
<tr>
<td>Mining, Utilities</td>
<td>.28</td>
<td>84%</td>
<td>.25</td>
<td>.09</td>
</tr>
<tr>
<td>Other Activities</td>
<td>.52</td>
<td>64%</td>
<td>.6</td>
<td>0</td>
</tr>
</tbody>
</table>

In Table 1.7, Other Activities and Manufacturing stand out for a few
reasons. Schooling seems to be important in explaining Manufacturing and Other Activities differences. 36% of the time the measure is negative, but it is quite high when positive. Figure 1.3 provides a visual of the distribution which appears vaguely normal with a mean above zero, but a large variance. The key finding is that schooling seems to be more important in Other Activities and Manufacturing relative to other industry pairs, but economies show a great deal of variance in the degree of importance of the schooling in value added.

Table 1.7: Baseline $m^2_{OA,man}$

<table>
<thead>
<tr>
<th>Industry</th>
<th>Mean $m^2_{j,man}$</th>
<th>Median</th>
<th>1st Quartile</th>
<th>3rd Quartile</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agriculture</td>
<td>.36</td>
<td>.13</td>
<td>.07</td>
<td>.26</td>
</tr>
<tr>
<td>Construction</td>
<td>-.69</td>
<td>.002</td>
<td>-.15</td>
<td>.16</td>
</tr>
<tr>
<td>Wholesale,Retail...</td>
<td>.12</td>
<td>.01</td>
<td>.02</td>
<td>.24</td>
</tr>
<tr>
<td>Transportation,...</td>
<td>.01</td>
<td>.02</td>
<td>-.05</td>
<td>.15</td>
</tr>
<tr>
<td>Mining, Utilities</td>
<td>.03</td>
<td>.02</td>
<td>.004</td>
<td>.06</td>
</tr>
<tr>
<td>Other Activities</td>
<td>.34</td>
<td>.23</td>
<td>-.19</td>
<td>.63</td>
</tr>
</tbody>
</table>

The final table, 1.8, shows the Alternative measure for Manufacturing and other industries. There remains very little role for schooling in explaining
productivity differences between Manufacturing, Construction, Wholesale, retail, etc. and Transportation and communication, but the results between Agriculture and Other Activities is magnified.

Table 1.8: Alternative $m^2_{j,man}$

<table>
<thead>
<tr>
<th></th>
<th>Mean $m^2_{j,man}$</th>
<th>Median</th>
<th>1st Quartile</th>
<th>3rd Quartile</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agriculture</td>
<td>.79</td>
<td>.31</td>
<td>.18</td>
<td>.63</td>
</tr>
<tr>
<td>Construction</td>
<td>-1.27</td>
<td>.034</td>
<td>-.27</td>
<td>.34</td>
</tr>
<tr>
<td>Wholesale,Retail...</td>
<td>.2</td>
<td>.027</td>
<td>-.04</td>
<td>.14</td>
</tr>
<tr>
<td>Transportation,...</td>
<td>.09</td>
<td>.05</td>
<td>-.1</td>
<td>.33</td>
</tr>
<tr>
<td>Mining, Utilities</td>
<td>.1</td>
<td>.05</td>
<td>.01</td>
<td>.15</td>
</tr>
<tr>
<td>Other Activities</td>
<td>1.03</td>
<td>.53</td>
<td>-.39</td>
<td>1.54</td>
</tr>
</tbody>
</table>

The data provided gives little notion of individual economies. In Figure 1.4 maps the efficiency units per worker and output per worker for African economies where all sectors are normalized to Agriculture. On the horizontal axis we note the large disparity of value added per worker and the relative similarity of education per worker.
Recall the initial composition of aggregate value added per worker.

\[ \frac{Y}{N} = \sum_{j=1}^{J} A_j \frac{L_j}{N_j} \frac{N_j}{N} \]

The counterfactuals will recompose the sectors in two ways. The first case is the adoption of Other activities schooling levels in all sectors. Since Other activities is the sector with the higher schooling this provides a reasonable number for education improvement in these economies. The second case is more conservative. It takes the agriculture efficiency units per worker \( \frac{L_{ag}}{N_{ag}} \) and replaces them with the manufacturing efficiency units per worker. This is a simple improvement of schooling in agriculture to the level of an "urban worker" defined as one who works in the manufacturing sector. The findings will be presented relative to aggregate \( \frac{Y}{N} \). Both counterfactuals are encouraging about the possibility of improvements from education, but do not address the huge aggregate differences in GDP between rich and poor economies. The economies are grouped by the most recent Latin America and African censuses, and the 10% of economies which have the largest agriculture sectors indicating less development.

<table>
<thead>
<tr>
<th></th>
<th>Case 1 mean</th>
<th>Case 1 median</th>
<th>Case 2 mean</th>
<th>Case 2 median</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Economies</td>
<td>1.24</td>
<td>1.2</td>
<td>1.04</td>
<td>1.03</td>
</tr>
<tr>
<td>Latin America</td>
<td>1.119</td>
<td>1.2</td>
<td>1.03</td>
<td>1.03</td>
</tr>
<tr>
<td>Africa</td>
<td>1.34</td>
<td>1.36</td>
<td>1.08</td>
<td>1.09</td>
</tr>
<tr>
<td>Poorest 10%</td>
<td>1.44</td>
<td>1.38</td>
<td>1.09</td>
<td>1.1</td>
</tr>
</tbody>
</table>

Composition plays an important role: economies with larger agriculture economies are more responsive to schooling changes in the aggregate. This is a very simple model, but it is informative about the extent to which schooling is important. We can perhaps increase the poorest economies output per worker by 44% by educating populations at the level of the most educated industries, or
increase GDP by 10% by educating rural populations to match the education of manufacturing workers. The second chapter provides a deeper analysis of these relationships.

1.5 Conclusion

Industries have significant differences in education and value added per worker. In two industries this paper finds that schooling can account for a significant share of the differences in output per worker. First, the Agriculture industry has significantly less schooling and value added per worker than other industries. While education is not the most important factor, between 15 and 30% of differences in industry value added could be explained by education in a median economy. “Other Activities” is an industry which includes a wide variety of high skill industries and is, not surprisingly, more highly educated than other industries. This industry tends to have higher value added per worker than other industries with some exceptions. While the variance of the share of other industries explained by schooling is quite high, an average value is around 20-30%.

These findings are interesting in the light of structural transformation, since we know Agriculture shrinks and Other activities increase as an economy develops. This chapter establishes the relationship between schooling and structural transformation which will be explored theoretically in the next chapter.
Low income economies typically have a large agriculture sector and few workers with schooling. Economic growth is accompanied by two trends in the labor force. First, the share of workers employed in the agriculture sector decreases significantly. Second, schooling goes from scarce to common. In Brazil, from 1960 to 2000, the share of workers with secondary education has increased from 3% to 25% while the share of workers in agriculture has decreased from 55% to 18%.

Agriculture sectors tend to have few workers with secondary educations and most workers with a secondary education work in other sectors, so structural transformation is movement of workers from a largely unschooled sector to sectors which employ greater shares of schooled workers. From the perspective of a farmer in a low income economy, education is a path to a high paying job in the non-agriculture sector. However, in many economies young farmers are not educating at extraordinary rates, which is evidence of high costs of education offsetting the benefits of higher income. Indeed, one can find many anecdotes of long journeys to poorly run schools. To address these concerns, policymakers in many economies design programs to increase access to schools and to improve the quality of education.

Taking these observations together, there is a narrative explaining structural transformation: schooling costs decrease as a result of policy and growth, so a greater share of the population is educated. This allows for a reallocation of workers to the higher skilled non-agriculture sectors. This mechanism has been used in the literature, most notably Casselli and Coleman (2001). If we employ policy that reduces the cost of education through increased efficiency and access, will this promote structural transformation? Is education an
effective policy channel to move an economy from the relative poverty of an agriculture economy to an “industrialized” economy? This paper asks to what extent changes in the costs of acquiring schooling drives structural transformation.

In order to explore the interaction of schooling and labor choices which bring about structural transformation, the paper builds on existing multi-sector growth models. It incorporates the major explanations of structural transformation during growth found in the literature, including a household with non-homothetic preferences as in Kongsamut et al (2002) and sectors with different rates of productivity growth following Ngai and Pissarides (2007). In addition to standard agriculture and non-agriculture sectors, an education sector is also modeled allowing for endogenous schooling costs from optimal labor allocations. Each sector, agriculture, non-agriculture and education, has a unique technology using capital, unschooled labor and schooled labor. A steady state equilibrium is calibrated to match a number of economies both initial and final observations of structural transformation. The schooling parameters in the calibration of the final economy are changed to the initial calibrations schooling parameters creating a counterfactual economy which experiences all changes but those in schooling technology. This counterfactual economy accounts for the structural transformation from increased efficiency and lower costs in schooling.

Important to this measurement is the role of schooled and unschooled labor in each sector. Agriculture and other sectors display significantly different patterns of schooling use. Agriculture sectors tend to employ fewer schooled workers, but as the number of schooled individuals in the economy increases the agriculture sector sees a larger increase in the share of schooled workers. Education sectors employ a large share of schooled workers even in economies where schooled labor is scarce. Since the starting point in the education sector is already high, the increase in the share of schooled workers is smaller when the
economy’s total share of schooled workers increases. The paper will consider sector differences among schooled and unschooled labor along two dimensions, intensity and elasticity. This captures the need for doctors or teachers in the non-agriculture or education sector, but also the relative difficulty of replacing these schooled workers.

The two varieties of labor are chosen optimally by the household in this dynamic setting. Because the education sector is subject to market wages and prices, the cost of schooling is endogenous. The education sector employs many teachers and as the share of schooled workers grows the cost of hiring teachers falls. Capital inputs are cheaper as the economy’s non-agriculture sector is more productive. As is normal in the human capital literature, there is some foregone labor time associated with educating. This time cost in a developing country can include schooling time, commuting, the time of parents and relatives providing school related assistance and transportation. The number of school locations, road quality, public transportation and the style or quality of teaching can all affect the amount of labor time foregone to provide an education. The counterfactuals consider both changes in the efficiency of the observed education sector and the decrease in foregone labor required to educate a student.

Data from the Integrated Public Use Microdata Series (IPUMS) International is used to discipline the sector technologies. The cross country calibration builds on both the structural transformation empirical literature including Valantinyi and Herrendorf (2008) and the empirical labor literature including Katz and Murphy (1993), but employs new data considering the relative elasticity and skill intensity of sectors. The initial and final steady state calibrations have similar technology and preference parameters, but differ in sector total factor productivity, foregone labor time for education, a demographic parameter which accounts for changes in the share of young workers and a discount factor which addresses
interest rate distortions from requiring steady state equilibrium. The individual economy calibration targets the share of schooled workers, the size of sectors, investment behavior and GDP per capita. Changes in the sector total factor productivity from the economies initial calibration to final calibration correspond to the existing theories of structural transformation: they increase, with agriculture increasing more than non-agriculture. One can turn off the alternative channel through education by considering the economy’s final calibration altered to take the education TFP and schooling parameter associated with the initial period. The paper also considers the long run implications of a particular improvement in schooling efficiency starting from an economy’s initial calibration.

The counterfactual experiments show most of the schooling levels are accounted for by changes in schooling parameters, but very little of the structural transformation is accounted for by schooling cost parameters. While schooling can account for a small share of structural transformation in some economies, the median difference between agricultural employment in the data and the counterfactual is near zero. The second counterfactual experiment where each economy’s initial calibrations are modified through improved efficiency in the education sector also shows little structural transformation. While the median share of workers with schooling increased by 102%, the share of works employed in agriculture decreases by only 2.7%. What technology would create large structural transformation from changes in schooling cost? If the elasticity of substitution of schooled and unschooled work is sufficiently inelastic, schooling costs account for a significant amount of structural transformation.

These findings have interesting implications for both those making policy recommendations to developing economies and growth researchers. This model finds that policies which improve the efficiency of education may improve the output of an economy, but do not lead to structural transformation. If economies
had not improved their schooling efficiency, non-agriculture sectors would remain large but employ fewer schooled workers. The increase in productivity of the agriculture sector remains the most important factor in explaining structural transformation. Human capital increases in the agriculture sector do not account for productivity increases known to occur in agriculture. Unless there is new evidence supporting a significantly lower elasticity of substitution of schooled and unschooled workers in developing economies, the schooling cost theory of structural transformation seems implausible. These findings pose an interesting puzzle for theories which account for most of income differences through human capital such as Manuelli and Seshadri (2007). Structural transformation is an important growth phenomenon which may be largely independent of observable human capital accumulation.

There has been some investigation of schooling composition in various sectors. Hendricks (2010) measures the amount of schooling which is accounted for by sector differences. The sector composition accounts for at most 20% of schooling differences cross countries. Hendricks aggregates his data from similar sources and his findings compliment the findings in this paper. The findings of the counterfactual show 20% of schooling changes accounted for in the median when schooling cost parameters are unchanged. Chapter 1 of this dissertation looks for sector specific worker productivity which can be attributed to schooling. Schooling accounted for more of the cross country non-agriculture productivity differences than cross country agriculture differences in worker efficiency.

explore how innate skill can account for the increases in agriculture efficiency during structural transformation. The most related papers are those of Caselli and Coleman (2001, 2006). There is a later section which responds to the 2001 paper on education and the convergence of the northern and southern United States. This paper could be viewed as an empirical exploration and model of the world technology frontier described in the 2006 paper of Caselli and Coleman.

The paper proceeds as follows: Section 2 establishes facts motivating the question. Section 3 presents the model. Section 4 provides background on data and the calibration of the model, and an examination of the fit. Section 5 presents findings, focusing on the two counterfactual exercises. The section finishes with a comparison to Caselli and Coleman (2001) through an alternative calibration closer to the technology used in that paper, which demonstrates the importance of the technology calibration. Section 6 concludes.

2.2 Motivation

Structural transformation drastically alters an economy in a single generation. During these transformations, we also see large changes in schooling measured as share of workers with secondary education. Figure 2.1 shows the structural transformation for eleven countries for periods of 10 to 40 years. The pattern is consistent: movement from high agriculture and low schooling in the lower right corner of the figure to low agriculture employment and high shares of workers schooled. The correlation of share of workers schooled and share of workers in agriculture is -.55 and the correlation of the log share of workers schooled and log share of workers in agriculture is -.69. These are not curious correlations, they are consistent with a common narrative of getting an education and moving off the family farm. A common narrative and an interesting correlation are motivation enough for serious modelling, but one would like further support of this narrative.

One requirement of a narrative where workers educate in order to leave
agriculture are clear differences between sector employment of schooled workers. The agriculture sector should be employing fewer schooled workers and the non-agriculture sector should employ a greater number of workers. This is visible in every country investigated in this paper. Figure 2.2 shows the share of workers with schooling in the agriculture and non-agriculture sectors for the same economies shown in Figure 2.1 with the addition of economies which do not have enough data to show long run structural transformation. For low education economies the difference is especially large, but it is also significant in countries with higher rates of schooling. In the United States the majority of agriculture workers since the 1980 census have secondary education, but the non-agriculture secondary completion rate is significantly higher. These findings are not unique to the secondary schooling definition of schooling: the pattern is similar for primary school completed and various aggregates including standard Mincer coefficient weightings as discussed in Schreck (2009).

This paper is also interested in the characteristics of the education sector.
Figure 2.2: Share of schooled workers in agriculture and non-agriculture sectors.

Figure 2.3 compares the share of schooled workers in the education and non-agriculture sector. The sector which is referred to as "non-agriculture" excludes the education sector as well as agriculture. Here the education sector has a greater share of schooled workers than the non-agriculture sectors. Again the biggest differences are among lower schooling economies. Agriculture shows large increases in schooled share after the non-agriculture share is large. The schooled share in education increases while non-agriculture schooling share is relatively low. If the non-agriculture sector has a 30% share of schooled workers, one would expect between 70% and 90% of workers in the education sector to be schooled. In contrast those economies show between 3% and 12% schooled workers in agriculture. While most economies in the sample have fewer than 40% of workers schooled in the non-agriculture sector, only four economies in the have fewer than 40% of its workers schooled in the education sector. The size of the education also varies, economies that school more students have a greater share of employment in the education sector. The correlation between share of workers
in the education sector and the share of workers schooled is .66. The size of the education sector and the quality of labor increase with the share of schooled workers in the economy. A model of the schooling choice and the education sector will provide further insight into these relations.

2.3 Model

An infinitely lived stand-in household has a measure 1 of labor each period. This household consumes two goods, agriculture and non-agriculture, and chooses capital and schooling each period. There are three sectors, non-agriculture, agriculture and education, each of which uses three inputs, capital, schooled labor and unschooled labor. The agriculture sector produces an agriculture consumption good. The non-agriculture sector produces consumption and capital goods. The education sector produces schooled workers. Along with education sector labor costs, the household also foregoes unskilled labor in order to school its workers. The household’s preferences are non-homothetic; there is a subsistence requirement for agriculture consumption. The consumption goods correspond to
value added measures in the data. The agricultural good is measured at the gate of the farm and any transportation and preparation services are considered non-agriculture goods.

The stand in household is endowed with a fixed measure of labor each period normalized to 1. A measure \( L_{s,t} \) of this labor is schooled. “Schooled” will refer to secondary school or 12 years of schooling completed. Secondary education is an internationally recognized schooling milestone which is highly correlated with other aggregate schooling metrics. Alternative measures are primary completion and university completion. Primary completion is not informative in high education economies where almost all workers have primary education. Share of workers with university completed is very low for low income economies and in these cases may not be as correlated with average years of schooling or share with primary education. At the initial period \( t = 1 \) the household is endowed with a schooled labor share \( L_{s,1} \). The household is also endowed with initial capital \( k_1 \). Before each period \( \delta \) share of laborers exit the labor force and measure \( \delta \) unskilled workers enter the labor force. This is a simple demographic structure intended to capture of the depreciation of schooled workers through retirement. Each period the household can educate \( s_t \in [0, \delta] \) of the new workers. Most education is received early in life, so this bounds the schooling choice function. The following law of motion will govern the share of schooled labor:

\[
L_{st+1} = s_t + (1 - \delta)L_{st}
\]

The household has preferences for two goods: an agriculture good and a non-agriculture good. These are not final goods, but instead value added shares of final goods. For example, a hamburger would consist of an agriculture share from the raising of the cows and growing of the wheat, but most of the hamburger consists of food services, transportation and retail. Most final goods in a developed
economy consist of a mix of agriculture and non-agriculture value added where the agriculture component is relatively small even in food goods. The household has a subsistence requirement, $\bar{c}_a$, for the agricultural good. Therefore, preferences will be non-homothetic. This is standard for the structural transformation literature and will produce some structural transformation as income increases. The utility function representing preferences will be a generalized Stone Geary utility function. The generalization will allow me to build on the findings of Buera and Kaboski (2009), and Herrendorf, Rogerson and Valentinyi (2009), which discuss the substitutability of agriculture and non-agriculture goods in the value added and final good characterizations. The functional form is as follows:

$$U\left(\{c_{at}, c_{nt}\}\right) = \sum_{t=0}^{\infty} \beta^t \left( \zeta \left(c_{at} - \bar{c}_a\right)^{\frac{\nu-1}{\nu}} + (1 - \zeta) \left(c_{nt}\right)^{\frac{\nu-1}{\nu}} \right)^{\frac{\nu(1-\rho)}{\nu-1}} - 1$$

Technologies for all three sectors are similarly modeled as constant elasticity of substitution production functions. Three inputs are used in production: schooled labor, unschooled labor and capital. Capital and labor will have a constant elasticity of substitution of one in all sectors, Cobb Douglas. Capital income shares will be different for different sectors. The elasticity of substitution of labor and weights on labor will be different in different sectors. It makes sense that different sectors provide different returns to education. Elasticity of substitution is less intuitive so I will provide an example. As the relative wage of schooled workers increases, farmers would be first to substitute from skilled workers to unskilled workers. Next industries which requires engineers to run machines may substitute, then the financial sector, and finally health care and education where education is most “necessary” for production. If this is true, we should see the smallest cross country and time series differences in schooling share in the education sector and the largest differences in the agriculture sector. I will show in
the calibration that this is exactly what we see in cross country data. We do not see all sector's schooling shares increasing at a similar rate as would be the case if the elasticity of substitution were the same. While the functional forms are simple, the differences in parameters provide richness to the model which allows the model to be successful in capturing sector education composition. Each sector has a productivity parameter which will capture the structural transformation theory of agricultural productivity outpacing non-agriculture productivity.

\[
Y_a = A_a k_a^{\theta_a} \left( \alpha_a l_{u,a}^{\sigma_a-1} + (1 - \alpha_a) l_{s,a}^{\sigma_a-1} \right)^{(1-\theta_a)\frac{\sigma_a}{\sigma_a-1}}
\]

\[
Y_n = A_n k_n^{\theta_n} \left( \alpha_n l_{u,n}^{\sigma_n-1} + (1 - \alpha_n) l_{s,n}^{\sigma_n-1} \right)^{(1-\theta_n)\frac{\sigma_n}{\sigma_n-1}}
\]

\[
s_t = \min \left\{ A_e k_e^{\gamma} \left( \alpha_e l_{u,e}^{\sigma_e-1} + (1 - \alpha_e) l_{s,e}^{\sigma_e-1} \right)^{(1-\gamma)\frac{\sigma_e}{\sigma_e-1}}, \frac{l_{u,f}}{\psi} \right\}
\]

All capital is created in the non-agriculture sector. While this is clearly an abstraction, it provides some interesting minor channels. There is significant evidence pointing to intermediate goods as an explanation productivity increases in agriculture. Capital is a significant share of agriculture spending; improvements in the productivity of non-agriculture will make capital abundant and agriculture labor more productive. If non-agriculture is relatively schooling intensive, an increase in schooling will also decrease the price of capital, increasing agriculture productivity. The same is of course true in schooling, although the effect is smaller since the capital share of schooling is relatively smaller. I do not differentiate from land and capital because land scarcity is not a large issue over long time horizons. Land is a relatively small share of capital as measured by Valentyni and Herrendorf (2008).

The schooling sector produces a schooling good. In addition to the schooling good, the households forgo a measure of unschooled time \( l_{f,t} \).
proportional to the share of schooling that period, $s_t$. This will be an abstraction for
students’ time and all other costs associated with the schooling of a young person.
Potential time costs include labor time spent in school for the pupil, time spent in
school for students who inevitably don’t complete their secondary education, time
spent commuting to and from school, time spent preparing for school, labor time of
family members spent assisting children in their schooling and transporting
students and any other time costs of other individuals for schooling which are not
captured within the formal education sector. There is also the possibility of time
subsidies if schooling frees parents to work in the labor force. One can imagine
that foregone labor parameter $\psi$ being determined by policy factors such as the
location of schools, the quality of roads, the time efficiency of the teaching style,
the efficiency of retaining students and public transportation. Obviously, $\psi$ may
also capture demographic and social trends, such as the employment of women.
Suffice it to say any aggregate abstraction of the schooling system requires
simplification from many real world complexities. The parameter $\psi$ will serve that
role within my model.

Equilibrium

This paper considers a dynamic competitive sequence of market equilibrium.
There are complete markets, so I can solve the social planner’s problem for the
allocation and then solve for prices. Prices are useful for the calibration,
competitive equilibrium is the needed equilibrium concept. The emphasis in this
paper will be on comparing steady states. There is ongoing work to extend the
equilibrium to model the entire transition.

A sequence of markets competitive equilibrium is a list of allocations
\{c_{a,t}, c_{n,t}, l_{s,t}, l_{f,t}, s_t, l_{u,a,t}, l_{s,a,t}, l_{u,n,t}, l_{s,e,t}, k_t, k_{a,t}, k_{n,t}, k_{e,t}\}and
prices\{p_{n,t}, p_{a,t}, p_{e,t}, w_{s,t}, w_{u,t}, r_t\} such that households maximize utility, firms
maximize profits and markets clear.
Household problem

The household takes \( p_{nt}, p_{at}, p_{st}, w_{st}, w_{at}, r_t \) as given, the household chooses \( c_{at}, c_{nt}, l_{at}, l_{ft}, k_t, s_t \) to solve the following:

\[
\max_{(c_{at}, c_{nt}, l_{at}, l_{ft}, k_t, s_t)} \left\{ \sum_{t=0}^{\infty} \beta^t \left( \zeta (c_{at} - \bar{c}_a)^{\frac{\nu-1}{\nu}} + (1 - \zeta) c_{nt}^{\frac{\nu-1}{\nu}} \right)^{\frac{\nu(1-\rho)}{\nu-1}} - 1 \right\}
\]

such that for each \( t \),

\[
p_{at} c_{at} + p_{st} s_t + p_{nt} (c_{nt} + k_{t+1} - (1 - \delta_k)k_t) \leq w_{st} l_{s,t} + w_{at} (1 - l_{s,t} - l_{ft}) + r_t k_t
\]

\[
l_{s,t+1} \leq s_t + (1 - \delta)l_{s,t}
\]

\[
s_t \leq \frac{l_{ft}}{\psi}
\]

\[
k_t \geq 0, s_t \in [0, \delta]
\]

The budget equation includes the purchase of the schooling good, agriculture goods, non-agriculture goods including capital. The rate of depreciation of capital is \( \delta_k \). The market labor which is schooled is denoted \( l_{s,t} \). Foregone market time required for schooling is \( l_{ft} \). The market labor which is unschooled is then \( 1 - l_{s,t} - \psi s_t \).

Agriculture Firms problem

Each period, taking \( p_{at}, w_{st}, w_{at} \) as given the agriculture firm chooses \( l_{u,a,t}, l_{s,a,t}, k_{a,t} \) so that

\[
\max_{\{l_{u,a,t}, l_{s,a,t}, k_{a,t}\}} p_{at} A_{at} k_{a,t}^{\theta_a} \left( \alpha_a l_{u,a,t}^{\frac{s_a-1}{s_a}} + (1 - \alpha_a) l_{s,a,t}^{\frac{s_a-1}{s_a}} \right)^{1-\theta_a} l_{u,a,t} - w_{st} l_{s,a,t} - w_{at} l_{u,a,t} - r_t k_{a,t}
\]

Non-agriculture Firms problem

Each period, taking \( p_{nt}, w_{st}, w_{at} \) as given the non-agriculture firm chooses \( l_{u,n,t}, l_{s,n,t}, k_{n,t} \) so that

\[
\max_{\{l_{u,n,t}, l_{s,n,t}, k_{n,t}\}} p_{nt} A_{nt} k_{n,t}^{\theta_n} \left( \alpha_n l_{u,n,t}^{\frac{s_n-1}{s_n}} + (1 - \alpha_n) l_{s,n,t}^{\frac{s_n-1}{s_n}} \right)^{1-\theta_n} l_{u,n,t} - w_{st} l_{s,n,t} - w_{at} l_{u,n,t} - r_t k_{n,t}
\]
Education Firms problem

Each period, taking \( p_{s,t}, p_{n,t}, w_{s,t}, w_{u,t} \) as given the education firm chooses \( l_{u,e,t}, l_{s,e,t}, c_{e,t} \) so that

\[
\max_{\{l_{u,e,t}, l_{s,e,t}, c_{e,t}\}} p_{s,t} A_{e,t} k_{e,t}^\gamma \left( \alpha_e l_{u,e,t}^{\sigma_e-1} - (1 - \alpha_e) l_{s,e,t}^{\sigma_e-1} \right) \left( 1 - \frac{\sigma_e}{\sigma_e-1} \right) - w_{s,t} l_{s,e,t} - w_{u,t} l_{u,e,t} - r_t k_{e,t}
\]

Market Clearing conditions

\[
l_{s,t} = l_{s,a,t} + l_{s,n,t} + l_{s,e,t}
\]

\[
1 - l_{s,t} - l_{f,t} = l_{u,a,t} + l_{u,n,t} + l_{u,e,t}
\]

\[
k_t = k_{a,t} + k_{n,t} + k_{e,t}
\]

\[
A_{a,t} l_{a,t}^{\theta_a} \left( \alpha_a l_{u,a,t}^{\sigma_a-1} + (1 - \alpha_a) l_{s,a,t}^{\sigma_a-1} \right) \left( 1 - \theta_a \right) \frac{\sigma_a}{\sigma_a-1} = c_{a,t}
\]

\[
A_{n,t} l_{n,t}^{\theta_n} \left( \alpha_n l_{u,n,t}^{\sigma_n-1} + (1 - \alpha_n) l_{s,n,t}^{\sigma_n-1} \right) \left( 1 - \theta_n \right) \frac{\sigma_n}{\sigma_n-1} = c_{n,t} + k_{t+1} - (1 - \delta_k) k_t
\]

\[
A_{e,t} l_{e,t}^{\gamma} \left( \alpha_e l_{u,e,t}^{\sigma_e-1} + (1 - \alpha_e) l_{s,e,t}^{\sigma_e-1} \right) \left( 1 - \gamma \right) \frac{\sigma_e}{\sigma_e-1} = s_t
\]

A steady state is an equilibrium where \( l_{s,t} = l_{s,t+1}, k_t = k_{t+1} \) for all \( t \). The steady state allows for some clear comparisons, but is an imperfect representation of structural transformation. The steady state requires an assumption that there is no capital or schooling growth at the point of calibration, something which is empirically false. It is likely that these economies are “in transition”, however transition to any steady state makes equally little sense since we can expect large growth in sector productivity which are different for different sectors. The notion of a balanced growth path is problematic since the growth of schooling is unbalanced because the share of schooling is bounded. It would be possible to consider transition to a balanced growth path, however choosing the appropriate balanced growth path becomes rather arbitrary. Therefore, I will present the most straightforward exercise which is common within the literature maintaining awareness of some shortcomings of this method.

35
2.4 Data and Calibration

Data

The primary data source for this paper is the IPUMS International collection of census data, from which 56 censuses from 21 countries are used. The data available from this source is broad and deep. IPUMS has done great work standardizing data so that countries are comparable. The size of the samples which consist of 5 to 10% of most populations allow for reliable subsamples for even the smallest sectors. The data ranges from high income to low income, from most regions of the world and years from 1960 to 2002. There is a strong representation of Latin American countries. IPUMS censuses have information on employment, sector, occupation, age, gender, home ownership and of course education. Most of the calibrations in this paper will focus on aggregates for sector and education. I will be using the Penn World Tables 6.3 for real GDP per capita and investment share data used for the calibration of growth and capital.

The definition of schooling is secondary education completed or 12 years of schooling. Data is also available for primary schooling and university completion, so there can be as many as 4 categories of schooling which are relatively consistent in all censuses. Since there are potential differences in the amount of primary schooling which is not captured by the secondary schooling coefficient, I create an adjusted schooling measure similar to that in Caselli and Coleman (2006). The key assumption is that no primary and primary schooling completion categories are perfect substitutes and secondary and university completion are perfect substitutes. Weights on the shares are applied using a Mincer coefficient of 10%, since country specific Mincer returns are not available for all economies and 10% is roughly the average and median of the Bils and Kelnor (2000) Mincer coefficients as well as a subsample corresponding to economies which are in both data sets. This allows the differences in the
unschooled and schooled to be associated with 12 years of schooling difference. I
do not want to capture the pure human capital effects. I renormalize the measures
so that the total measure of labor is 1. The simple secondary shares and the
renormalized weighted shares for the four sectors are very similar. Renormalizing
is necessary to keep human capital choice limited to a single dimension. Because
the productivity increase associated with no primary to primary and secondary to
university are relatively small, this should cause small biases towards higher
education costs. A potential future extension could address ways to include a
general human capital increase along with the discrete schooling achievement.

**Calibration**

The model has complete markets, so it can be solved using the social planners
problem, however prices are useful in the calibration of relative wages. The
approach taken solves a social planner’s problem for allocations, than solve for
prices. Since the commodity space is convex and functions are concave, the
steady state exists for $A_a$ large enough for subsistence. Corner solutions,
specifically $L_{st} = 0$, are common if schooling cost parameter $\psi$ is too great or the
schooling sector $A_e$ is too low. While there are no Inada conditions keeping the
steady state interior, the interesting observed cases are where schooling levels are
low but not zero.

**Endowments**

The period length will be 5 years time. The longer time period is appropriate for
the long run nature of decisions and the frequency of census data. Since the
model assumes constant population size, the demographic parameter is $\delta$, which
is the number of older workers leaving the labor force and the number of workers
entering the labor force. The value is chosen to be consistent to the share of
young employed workers ages 20-25 and 30-35 as a share of the entire labor
population. This is calculated uniquely for each economy and takes values between .1 and .15.

Preferences

There are four major preference parameters to calibrate: $\zeta$ the weight on agricultural goods, $\bar{c}_a$ the subsistence level in agriculture, $\upsilon$ the elasticity-like parameter for various consumption types and $\rho$ the intertemporal elasticity of substitution. Good data on relative prices of value added agriculture and non-agriculture goods is not available for the countries for which there is census data. Therefore, this portion of the calibration relies on literature findings for these parameters. $\zeta$ is chosen to match the share of value added agricultural consumption when income is very high and subsidence does not impact choice. Since the United States has a level agriculture consumption share of near 1.5%, $\zeta$ is chosen to be below that level at $\zeta = .01$. It is difficult to jointly calibrate $\bar{c}_a$ and $A_a$ without an agriculture price, therefore $\bar{c}_a$ will be chosen to match findings from micro studies in India. Rosenweig and Wolpin (1993) and Atkeson and Ogaki (1996) use rural household data to estimate subsistence consumption in similar models finding 34% of average income in 1984 India. The $\upsilon$ is chosen to be close to Leontief as is found in Herrendorf, Rogerson and Valentinyi (2009).

Preferences describing the United States structural transformation using value added measurements best fit for low elasticity parameters, therefore $\upsilon = .1$ is chosen as a very inelastic parameter. The value $\rho$ takes a standard literature value 2, but is not relevant in the steady state calibration.

Technology

There are twelve parameters to calibration, four for each technology. The productivity parameters $A_a, A_n, A_e$ will be calibrated jointly for each economy, which will be discussed in a later section. Capital shares are taken from the literature. Labor parameters are taken from my own observations. Valentiynyi and
Herrendorf (2008) provide values for capital income share in the agriculture and non-agriculture $\theta_a = .5, \theta_n = .33$. This is consistent with the literature showing a larger capital share in agriculture as compared to non-agriculture. The labor income share in the education is about 90% of education value added, so the capital share is a low $\gamma = .1$. This can be verified for the United States using the BEA industry value added components: employee compensation shares in education are between 87% and 90% for 1998 and 2000. This is also a typical value used in the literature including papers like Restuccia and Vandenbrooke (2008).

Measuring elasticities of substitution in production functions is typically done by measuring the change of schooling levels for a given change in wage premium. This approach requires good data on wages with some heterogeneity such as region or time. Because wage data is not available in many developing countries and the project focuses on agriculture economies not best represented by the United States, the approach in this paper will be different. Instead of looking at how sectors respond to a wage change, the alternative is to consider the relative growth of sectors schooling share from a change in schooling. If a 1% increase in schooled workers leads to a 1% increase in the sector shares of schooled workers in all sectors, than we would expect sector elasticities to be about constant. If one sector saw a .5% increase and another saw a 2% increase, we would consider the second sector to be relatively more elastic. Since data is available on the relative data schooling sector shares for all of the economies, the relative sector elasticities can be measured. This is not intended to be the final word on sector elasticities, further scrutiny using rigorous estimations is desired. This paper can support differences in sector elasticities of substitution of schooled and unschooled workers. Looking at relative elasticities requires an absolute elasticity, so the elasticity of non-agriculture will be $\sigma_n = 1.4$ which is the Katz and Murphy
measurement of the elasticity of substitution of secondary schooled and unschooled workers in the United States. This is the same value used by Caselli and Coleman (2006). Since the United States has been a non-agriculture economy, this is a reasonable starting value. The findings section explores alternative calibrations of $\sigma_n$.

The measurement depends on the model assumption of a single wage for the schooled and unschooled in all sectors or long run labor mobility. The first order conditions provide the following relation between wage premium and the relative intensity of sector schooling. The intensity of schooled workers is inversely related with the wage premium and the weight on unschooled $\alpha_n$.

\[
1 - \frac{\alpha_n}{\alpha_n} \left( \frac{l_{s,n,i}}{l_{u,n,i}} \right) = \frac{w_{s,i}}{w_{u,i}}
\]

\[
1 - \frac{\alpha_a}{\alpha_a} \left( \frac{l_{s,a,i}}{l_{u,a,i}} \right) = \frac{w_{s,i}}{w_{u,i}}
\]

\[
1 - \frac{\alpha_e}{\alpha_e} \left( \frac{l_{s,e,i}}{l_{u,e,i}} \right) = \frac{w_{s,i}}{w_{u,i}}
\]

Since there are economies with the same wage, we should observe the following relations in the data:

\[
\log \left( \frac{1 - \alpha_n}{\alpha_n} \right) - \frac{1}{\sigma_n} \log \left( \frac{l_{s,n,i}}{l_{u,n,i}} \right) = \log \left( \frac{1 - \alpha_a}{\alpha_a} \right) - \frac{1}{\sigma_a} \log \left( \frac{l_{s,a,i}}{l_{u,a,i}} \right)
\]

\[
\log \left( \frac{1 - \alpha_n}{\alpha_n} \right) - \frac{1}{\sigma_n} \log \left( \frac{l_{s,n,i}}{l_{u,n,i}} \right) = \log \left( \frac{1 - \alpha_e}{\alpha_e} \right) - \frac{1}{\sigma_e} \log \left( \frac{l_{s,e,i}}{l_{u,e,i}} \right)
\]

When plotted in figures 4 and 5, the linear relationship from $\frac{l_{s,a,i}}{l_{u,a,i}}$ to $\frac{l_{s,e,i}}{l_{u,e,i}}$ seems evident. An OLS fit is good and $\frac{\sigma_a}{\sigma_n}$ and $\frac{\sigma_e}{\sigma_n}$ are significantly different from 1. The OLS regression is treating each individual time and country as an economy, using 56 members for the sample from 21 countries. This is reasonable because the time differences are large enough between economies that the structures are significantly different. The two expected trends of higher elasticity of substitution in agriculture and lower elasticity of substitution in education are
Figure 2.4: Non-agriculture and agriculture log schooling shares for sample countries.

supported by the measurements from the OLS regression. If the elasticity of substitution for non-agriculture is 1.4 the elasticity for agriculture is measured as \( \sigma_a = 1.8 \) so as schooling increases, the schooling intensity of agriculture increases relatively quicker than non-agriculture. Similarly, the elasticity of substitution of schooled and unschooled labor for education is measured as \( \sigma_e = 1.1 \). I also measured these parameters using the larger data set described in Chapter 1 and found similar results.

The most obvious trend in the figures is the highest level of schooling intensity, \( \ell_s \), in the education sector followed by non-agriculture then agriculture. If one were to restrict elasticity to 1.4 or Cobb-Douglas, there would be larger differences in the weights on unschooled \( \alpha_a, \alpha_n, \alpha_e \). The weights will follow the order \( \alpha_a < \alpha_n < \alpha_e \), and ratios can be calculated from the constant:

\[
\text{OLS coefficient} = -\sigma_n \log \frac{\alpha_n (1 - \alpha_a)}{\alpha_a (1 - \alpha_n)}
\]

With the relative levels of the weight parameters given by the OLS coefficient and elasticity, an absolute level must be determined from the wage premium. Using the Bils and Klenow (2000), I take the average wage premium for 12 years of
schooling and use this to calibrate the mean \( \frac{w_{s,t}}{w_{u,t}} \). I verify that the Bils and Klenow sample is representative of my sample by finding as many Mincer regressions that correspond to 24 of the censuses and measure the average from that census. This provides a base \( \alpha_n \), from which \( \alpha_a \), \( \alpha_e \) are calculated:

\[
\log \frac{1 - \alpha_n}{\alpha_n} = \frac{1}{\sigma_n} \sum_{i=1}^{I} \log \left( \frac{l_{s,n,i}}{l_{u,n,i}} \right) = \log \frac{w_s}{w_u}
\]

The resulting parameters are \( \alpha_n = .43, \alpha_a = .657, \alpha_e = .123 \). These values make sense for developing countries, but are too high for the United States. There is a large degree of sector specific technological progress made going from agriculture to non-agriculture, however this is not sufficient to explain the skill bias technology in the United States since 1960. Other papers have tackled skill biased technological progress in the United States using more sectors, but the focus of this paper is on the developing nations.

**Economy Calibration**

After all of the common variables have been calibrated, 5 parameters remain: \( A_n, A_a, A_e, \psi, \beta \) which will be calibrated to economy specific moments. These
parameters will be jointly calibrated to match share of workers schooled, share of workers in agriculture, share of workers in education, real GDP per capita and investment share of GDP. The real GDP per capita and investment share of GDP are from the Penn World Tables 6.3. Included are all economies where two censuses exist at least 9 years apart which display structural transformation. A steady state is calibrated for the earliest and most recent censuses available. Targetting the schooled share and the sector shares of workers is natural for this exercise. Investment ratios are important because investing in capital and educating workers are substitutes from the perspective of the households. The real interest rates are potentially lower than the interest rate in the United States in these periods. While the model is in steadystate, capturing the intertemporal decision through observed capital behavior is helpful. The real GDP per capita is important because the TFP level will effect schooling costs endogenously and the level of output relative to subsistence needs will effect households decisions. Data labor shares are all adjusted to exlude \( l_f \) so that the measure of labor in the data corresponds to what is measured in the model.

The calibration starts with finding sector labor shares, schooled and unschooled, which correspond with the aggregate schooling shares and technology parameters \( \alpha_n, \alpha_a, \alpha_e, \sigma_n, \sigma_a, \sigma_e \). The level of capital and sector capital shares are computed from the labor shares, real GDP per capita and investment to GDP ratio. The country specific discount factor \( \beta \) follows from the implicit real interest rate from the capital choice. The \( A_e \) follows from the steady state \( s = \delta L_s \) and the education labor and capital levels. Similarly the value of \( A_n \) and \( c_n \) follow from real GDP per capita, investment shares and capital. Given consumption preferences the value of \( c_a \) relative to \( c_n \) is determined allowing for the measurement of \( A_a \). The \( \psi \) parameter equates marginal cost of schooling and marginal return. \( \psi \) must then be used to correct labor shares and the process is repeated until \( \psi \)
These conditions are derived from the measurement of real $Y_t$.

$$\frac{Y_t}{p_{nt}} = (c_{nt} + k_{t+1} - (1 - \delta_k)k_t) + \frac{p_{nt}}{p_{nt}} c_{at} + \frac{p_{st}}{p_{nt}} s_t$$

While this does not correspond perfectly with the PPP measures used by the Penn World Tables, this abstraction is precise enough to capture the large changes in real GDP over long time periods: a magnitude of 10 for countries differing by 40 years.

**Calibration Example: Brazil**

Using an example economy, one can observe properties of the model and fit. First, the model does a good job fitting the sector shares of schooled and unschooled which are only indirectly calibrated through the regression. One can see that initially in 1960 there is very few schooled workers, most of whom are in non-agriculture. In the final period, 2000, the education sector is employing a larger number of schooled workers and the number of non-agriculture schooled workers in model and data closely match.
Table 1 shows how the model of cost of schooling per capita and cost per student in real GDP per capita units. While the expenditure on education has increased over this period of time, the cost per student has decreased dramatically. Most of this cost is hiring the skilled workers. Model wage premiums fall from $\frac{w_s}{w_u} = 10$ to $\frac{w_s}{w_u} = 2.65$, so the cost of teachers fell drastically as did the returns of gaining an education. While a return to secondary school of 10 may seem high especially for the United States, measurements of wage premiums in Brazil in 1970 support wage premia this high.

\[
\frac{\text{Education Cost}}{\text{real GDP per capita}} = \frac{p_s s_t + w_u \psi_s t}{Y_t}
\]

Table 2.1: Schooling Model for Brazil

<table>
<thead>
<tr>
<th></th>
<th>1960 Model</th>
<th>2000 Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Education Cost</td>
<td>4.55%</td>
<td>6.73%</td>
</tr>
<tr>
<td>real GDP per capita</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cost per student in real GDP per capita</td>
<td>11.375</td>
<td>2.1202</td>
</tr>
<tr>
<td>Share of cost hiring skilled workers</td>
<td>67.91%</td>
<td>80.78%</td>
</tr>
</tbody>
</table>

There are large decreases in cost, large increases in non-agriculture schooled workers and large increases in income in Brazil. One could view this as evidence supporting schooling costs as important for structural transformation. Figure 3 indicates that something else is likely more important: the share of unschooled workers in non-agriculture has also increased and the number of unschooled agriculture workers has decreased significantly. The counterfactual exercises will formalize this observation.

2.5 Findings

To evaluate the effect of schooling parameters two counterfactual parameterizations will be considered for each economy where there are two data points displaying structural transformation at least 10 years apart. The first counterfactual exercise considers structural transformation without changes in
schooling parameters. The schooling parameters will correspond with the initial steady state and all other parameters will correspond with the final steady state. The second counterfactual exercise demonstrate the introduction schooling parameters in lower income economies. Schooling parameters will change for each economy while other parameters will remain at the initial steady state. Both exercises find very little structural transformation, despite large changes in schooling.

Counterfactual Exercise 1

To evaluate the degree to which structural transformation occured due to schooling, the schooling parameters $A_{e,i}$, $\psi_i$ are set to the initial steady state level, while the other parameters $A_{n,i}$, $A_{a,i}$, $\beta_i$, $\delta_i$ are set to the final steady state for each economy. Relative to the change in schooled labor share, agricultural labor share and output which we observe in the data, how much change does one observe without changes in schooling costs? The findings are displayed in Figure 2.7. The median increase in schooled labor share is 20% of the schooling change observed. The schooling parameters are important for explaining schooling levels in all of the economies. The median increase in agricultural labor share is 100% which can be explained without changes in schooling parameters. While there is some effect in specific cases, there is no systematic effect of schooling parameters on the allocation of resources across sectors. The median output increase is 65% and varies significantly country to country. Schooling cost parameters are effecting output, but not through resource allocation.

Counterfactual Exercise 2

An alternative exercise is to consider the possible effects of policy through schooling cost parameters in a lower income economy. The parameter change will be typical for cost changes during structural transformation: a 50% increase in efficiency of the schooling sector, $A_{e,i}$, and a 50% decrease in the foregone labor
cost parameter, $\psi_i$. If these policies promote structural transformation, one should observe decreases in the agricultural labor share. Intuitively, this is like teachers training which allows teachers in a school to teach and graduate 50% more students and reducing students time cost by half through better school access. The median increase in schooled labor share is 102% demonstrating large effects from parameter changes. The median change in agriculture labor share is -2.7%. While some structural transformation effect is visible, it is minimal. The median output increase is 26%.

**Analysis**

Since schooling parameters have little effect on structural transformation, a natural question is what is causing structural transformation in these economies. Agriculture labor share falls from 55% to 18% in Brazil and from 40% to 15% in Greece, so one observes drastic structural transformation. Increases in the relative productivity of agriculture are observed in all the economies with structural
transformation. Figure 2.9 shows the relation between the difference in the size of agriculture sectors from the initial period to the final period and the change in the relative TFP $\frac{A_u}{A_e}$ in economies from the initial period to the final period. The economies which display large changes in agriculture labor shares tend to display large changes in relative TFP; the dramatic cases being Greece and Portugal. Economies which showed little change in structure also saw relatively low TFP changes. Economies which are initially closer to the subsistence requirement will have greater structural transformation resulting from non-homoethetic preferences. Brazil and Ecuador are good examples of a country with significant structural transformation due to initial low incomes.

One may think that an economy’s initial conditions will effect the economy’s response to changes in schooling parameter changes. An economy which is consuming agriculture for subsistence purposes may be more responsive to income increases related to more schooling, leading to more structural
Figure 2.9: Changes in relative productivity of agriculture and structural transformation. Figure 2.9 provides information on how structural transformation seen in Counterfactual Exercise 2 correspond to the subsistence level. Brazil and Ecuador have important subsistence requirements in the initial period and the model shows the most structural transformation as a result of this policy. While this trend holds at the extreme there are other factors like initial schooling costs which influence the change in labor share due to a 50% increase in efficiency and 50% decrease in foregone labor. A very low income economy in Africa may see some structural transformation resulting from more efficient schooling policy, however the largest decrease observed is Brazil at 8% so even the best projections from larger parameter changes are small. A policy maker wishing to industrialize would be better off attempting to improve agricultural productivity directly.
Comparison to Caselli and Coleman

In their 2001 paper, Caselli and Coleman use a model of education and structural transformation to explain the convergence of the North and the South in the United States. The model uses an abstraction called "skill" and simple education costs, which are similar to schooling and the education sector in the framework of this paper. The Caselli and Coleman paper have an unskilled agriculture technology and a skilled non-agriculture technology, so only unskilled workers are employed in agriculture and one where only skilled workers are employed in non-agriculture. The model presented in this paper is designed to address a different set of facts yet it is interesting to consider what is most important in driving the differences in findings. Instead of imposing extreme assumptions on the technology, this comparison will make use of a single variable in the calibration: the elasticity of substitution of skilled and unskilled labor. One interpretation of the Caselli and Coleman framework is that the substitution of skilled and unskilled labor is zero.
The literature finds the elasticity of substitution to be between 1 and 2, which is the basis for the calibration of non-agriculture elasticity of substitution. If this elasticity of substitution is lowered to $\sigma_n = .5$ the counterfactual economies see significantly more structural transformation as a result of schooling parameter changes. This calibration also shows significantly larger changes in the $\psi$ parameter, which is similar to the large changes in education costs required in the Caselli and Coleman framework.

The entire technology is recalibrated for elasticity of substitution of schooling $\sigma_n = \{.5, 1, 1.4, 2\}$ this captures the range of measured values of elasticity of substitution and a value which is below most estimates. For the $\sigma_n = 1$ the value of $\sigma_a$ will take the value of the 99% confidence interval to maximize the difference between elasticities of the two sectors. The counterfactual experiments are performed for each $\sigma_n$. Displayed in table 2.2 are median values for the structural transformation. In both experiments it is clear that there is more structural transformation which can be attributed to schooling parameter changes when the elasticity of substitution of schooled labor is low. Figure 2.11 provides the country by country findings which show changes resulting from improved schooling efficiency have much larger effects on agriculture labor share. The exception is the Philippines where a reasonable calibration could not be found for this counterfactual experiment. This is because the Philippines is a relative outlier where schooling is already high while a large share of workers is employed in agriculture. The foregone labor parameter $\psi$ was an order of magnitude higher in the inelastic case, giving values which seem implausible. With the elasticity $\sigma_n = 2$ there is less change in $\psi$ and more importance of the technology weights $\alpha_i$. This paper does not target wage premiums beyond the mean, however the higher $\sigma_n$ provides tighter distribution of wage premiums similar to that seen in the data.

Casselli and Coleman’s framework makes sense in a world where there is
Table 2.2: Findings with different elasticities

<table>
<thead>
<tr>
<th></th>
<th>$\sigma_n = .5$</th>
<th>$\sigma_n = 1$</th>
<th>$\sigma_n = 1.4$</th>
<th>$\sigma_n = 2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hypothetical</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median Agriculture change accounted for</td>
<td>50%</td>
<td>93%</td>
<td>100%</td>
<td>104%</td>
</tr>
<tr>
<td>Counterfactual Exercise 2:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median Agriculture Share change</td>
<td>-12%</td>
<td>-4.6%</td>
<td>-2.7%</td>
<td>-2.1%</td>
</tr>
</tbody>
</table>

Figure 2.11: Findings for Counterfacutal Exercise 2 for different technology parameterizations: Change of initial economies from a change of +50% TFP education and -50% $\psi$

significantly less substitution of skilled and unskilled workers. By taking their framework and calibrating it to be consistent to the literatures findings on the substitution of labor, the underlying mechanism no longer functions. Even at the most generous calibration supported empirically leaves little role for schooling in structural transformation. While the non-agriculture sector is less elastic than the agriculture sector, the difference is not large enough to demonstrate populations "educating" their way into non-agriculture.
2.6 Conclusion

If schooling productivity and costs remain unchanged during growth one should expect very little difference in the structural transformation. Schooled workers would continue to work in the non-agriculture sector, but so would large numbers of unschooled workers. This counterfactual non-agriculture sector would be less productive, but productivity increases in the agriculture sector would lead workers to the increased employment in non-agriculture observed in the data. This model is one of optimal decisions where there is not a role for government, but it does provide some evidence that educating a population to promote "industrialization" is not effective. Governments wishing to decrease the employment of the agriculture sector should focus on agricultural productivity. This paper confirms the common finding in the human capital literature that increases in education reduce the cost of education. The ability to observe this concept through the use of cross country sector data is a contribution. The idea that agricultural productivity growth quickly outpaces schooling growth in the agriculture sector is problematic for theories that want to account for most of growth through human capital improvements. The sector which shows the most productivity growth is also the sector which has the least schooling.

It would be interesting to extend this model to an equilibrium which captures the transition path. While this may provide better insight into the speed of transitions, it is unlikely to change the levels that result of varying policies. Further extensions should include a better model of intermediates in agriculture and how the increase in non-agriculture productivity from schooling effects the productivity of the agriculture sector.


