THREE ESSAYS IN QUANTITATIVE LABOR ECONOMICS

A Dissertation
Submitted to the Faculty of the
Graduate School of Arts and Sciences
of Georgetown University
in partial fulfillment of the requirements for the
degree of
Doctor of Philosophy
in Economics

by

Beom Sock Park, M.S.

Washington, DC
June 28, 2010
Copyright 2010 by Beom Sock Park  
All rights reserved
THREE ESSAYS IN QUANTITATIVE LABOR ECONOMICS

Beom Sock Park, M.S.

Thesis Co-Advisors: James W. Albrecht, Ph.D., Catalina Gutierrez, Ph.D. and Susan Vroman, Ph.D.

ABSTRACT

In chapter 1, we construct a multi-sector labor search and matching model to investigate how economic shocks affect labor markets in developing countries. Workers in our model are heterogeneous in their productivity, and they can be employed either in high or low productivity urban jobs or in agriculture. In urban labor markets, job search frictions exist. Economic shocks can destroy jobs, and workers become unemployed when their job matches are dissolved. Identical workers in our model can be employed in different sectors with different earnings. We calibrate our model to Nicaragua's labor market and then simulate impacts of financial shocks on wages and employment shares. We find that the economic shocks tend to have modest impacts on total employment but generate significant relocations of workers across sectors.

We validate the multi-sector model in chapter 2. Our results suggest that the model closely matches an average skill distribution across sectors in Indonesia in 1997, and that high-skilled workers are found in the productive formal sector, whereas low-skilled workers are located in production in either the agricultural or the informal sector.
In chapter 3, we use the theory of human capital investment developed by Becker and Tomes (1979) to quantify the impacts of human capital investment and the transmission of learning ability on intergenerational income mobility. Altruistic parents invest part of their income in their children's human capital accumulation. One's learning ability is formed partly by the ability of one's parents and partly by a random environment effect. We parameterize our model for 10 OECD countries and find that countries with high intergenerational income correlation tend to show high returns on per-unit human capital investment, while the process of learning ability transmission is mainly responsible for cross-sectional income inequality. Furthermore, we find that human capital subsidies can play a significant role in improving mobility. This mobility gain is obtained through higher returns from human capital investment among the poor.
# Table of Contents

Introduction........................................................................................................................................ 1

Chapter 1: The Impact of Economic Shocks on a Multi-sector Labor Market; Application to Nicaragua ................................................................. 5

1.1 Introduction..................................................................................................................................... 5

1.2 Model................................................................................................................................................ 10

1.2.1 Environment .................................................................................................................................. 11

1.2.2 Workers ......................................................................................................................................... 14

1.2.3 Formal Sector Firms ....................................................................................................................... 15

1.2.4 Wage Bargaining ............................................................................................................................ 17

1.2.5 Reservation Productivities ........................................................................................................... 18

1.2.6 Cutoff Productivities and Unemployment Values ........................................................................ 18

1.2.7 Steady-State Conditions in Urban Areas ...................................................................................... 21

1.2.8 No-Migration Condition(s) .......................................................................................................... 23

Unilateral (Rural-to-Urban) Migration ............................................................................................... 24
Chapter 2: Validating a Multi-sector Labor Search and Matching Model for Developing Countries

2.1 Introduction

2.2 Model
2.3 Data and Target Statistics................................................................. 45

2.3.1 Data Sources: Sarkernas vs. IFLS.......................................................45

2.3.2 Labor Composition and Labor Income.............................................47

2.4 Benchmark Parameterization.......................................................... 48

2.4.1 Implications of Benchmark Parameterization.................................51

2.5 Outcome Validation........................................................................... 52

2.5.1 Skill Distribution.............................................................................53

2.5.2 Wage Distribution...........................................................................54

2.5.3 Structure of Segmentation by Skill................................................55

2.5.4 Employed Skill Distribution..........................................................56

2.5.5 Unemployment Rates by Skill.........................................................57

2.5.6 Summary.........................................................................................58

2.6 Conclusion......................................................................................... 59

Chapter 3: Quantitative Study of Cross-Country Intergenerational Mobility….. 61
Table.......................................................................................................................... 99

Figure.......................................................................................................................... 108
Introduction

As many economists have noted, economic crises are recurrent phenomena. Between 1970 and 2008, there were 124 systemic banking crises, 208 currency crises, 63 sovereign debt crises, two oil shocks in the ’70s and the food and energy price shock in 2007-2008 (Verick et al., 2010). As economies become more integrated across borders, countries become more vulnerable to economic shocks. What are the welfare consequences of the shocks in developing countries? In order to answer this question, we should notice that since financial and insurance markets are not well developed in developing countries, labor stock or human capital is the primary household asset. Therefore, it is important to study the impact on labor markets to assess the welfare consequences of economic shocks in developing countries.

Labor markets in developing countries showed different adjustment processes in previous economic crises. For example, the average wage dropped by over 40% in countries like Mexico and Russia, and the average wage in Romania fell by 28%. Bulgaria and Chile had relatively more rigid wage settlements, so labor markets were adjusted through a drop in employment; 14% in Bulgaria in 1991 and 11% in Chile in 1982. These different adjustment processes reflect either the different institutional settings or the different nature of economic shocks that these countries faced.
For the purposes of policy intervention and impact evaluation regarding labor markets, we need to have some framework that coherently explains these changes after taking into account the different institutional settings or the nature of economic shocks. In our opinion, both neo-classical competitive framework and dual labor market models may not be well suited to describe the nature of labor market in developing countries.\(^1\) To make the list short, the competitive framework in which all workers are paid according to their labor productivity, does not fit the different sectoral wage determining mechanisms found in developing countries while we do not have any conclusive empirical finding to support labor dualism in developing countries.

Although unemployment is one of the main interests during the periods of economic crises, the magnitude of its changes is relatively small in developing countries. Instead, inter-sectoral labor relocation is commonly observed as well as urban-to-rural return migration. This observation further suggests that we need to deviate from competitive framework in analyzing the impact of economic shocks on labor markets in developing countries.

Economists have paid attention to structural inequality caused by intergenerational human capital transmission. If skill biased technical change favors those from income rich families since children from income rich families have more and better opportunities to acquire human capital than those from income poor families, this structural inequality reflects not only social injustice, but also economic inefficiency. A society where one’s economic outcome significantly depends on one’s family background cannot be called a fair society because one can be discriminated against in the race of economic success, due to his or her humble family background. If one cannot be given opportunities to fulfill one’s potential just because one was born in a poor region or country,

\(^1\)We will explain this point in detail when we motivate our model.
that is clearly economically inefficient.

Therefore, it is important, first, to know channels through which intergenerational human capital transmission is made. Second, we need to examine quantitatively how human capital investment affects children’s economic outcomes, and how this investment affects intergenerational income correlations. The answers to these questions may give some implications to human capital policy.

In chapter 1, we propose a multi-sector labor search and matching model. Workers in the model are heterogeneous in their productivity, and they can be employed either in high or low productivity urban jobs or in agriculture. In urban labor markets, job search frictions exist, in that it takes time for unemployed workers to find jobs and for firms to fill vacancies. Economic shocks can destroy jobs, and workers become unemployed when their job matches are dissolved. Identical workers in our model can be employed in different sectors with different earnings levels. We calibrate our model to Nicaragua’s labor market in the year 2001 and then simulate impacts of the financial shocks on wages and employment shares. We find that the economic shocks tend to have modest impacts on total employment but generate significant relocations of workers across sectors.

We validate the multi-sector model in chapter 2. We compare model outcomes with Indonesian labor markets in 1997. In our validation exercises, we find that our model closely matches an average skill distribution across sectors in Indonesia, and that high-skilled workers are found in the productive formal sector, whereas low-skilled workers are located in either the agricultural or the informal sector. Additionally, unemployment share increases with skill, which is also observed in the data. However, our model fails to match low unemployment shares for high skilled workers.
Overall, the model is able to explain the average behavior of workers in terms of employment composition and wages across sectors.

In chapter 3, we use the theory of human capital investment developed by Becker and Tomes (1979) to quantify the impacts of human capital investment and the transmission of learning ability on intergenerational income mobility. Altruistic parents invest part of their income into their children’s human capital accumulation. One’s learning ability is formed partly by the ability of one’s parents and partly by a random environmental effect. We parameterize our model for 10 OECD countries and find that countries with high intergenerational income correlation (or elasticity) tend to show high returns on per-unit human capital investment, while the process of learning ability transmission is mainly responsible for cross-sectional income inequality. We further study how the random component in learning ability transmission and human capital subsidies affect mobility. Although the degree of randomness is quantitatively negligible, human capital subsidies can play a significant role in improving mobility. A subsidy of 10% of mean life-time income reduces the intergenerational income correlation from 0.51 to 0.2. This mobility gain is obtained through higher returns from human capital investment among the poor.
Chapter 1

The Impact of Economic Shocks on a Multi-sector Labor Market; Application to Nicaragua

1.1 Introduction

As many economists have noted, economic crises are recurrent phenomena. Between 1970 and 2008, there were 124 systemic banking crises, 208 currency crises, 63 sovereign debt crises, two oil shocks in the ’70s and the food and energy price shock in 2007-2008 (Verick et al., 2010). As economies become more integrated across borders, the impact of the economic shock in one country or region tends to spread out to other countries or regions. How does the economic shock affect developing countries? Since financial and insurance markets are not well developed in developing
countries, labor stock or human capital is the primary household asset. Therefore, it is important to study the impact on labor markets to assess the welfare consequences of economic shocks in developing countries.

Employment and earnings are widely recognized as important channels through which economic shocks and policy reforms affect GDP growth and household welfare. However, it is unclear how unanticipated economic shocks affect labor market variables. For example, in countries like Mexico and Russia, average wages dropped by over 40% in previous crises, and the average wage in Romania fell by 28%. Bulgaria and Chile had relatively more rigid wage settlements, and so labor markets were adjusted through a drop in employment: 14% in Bulgaria in 1991 and 11% in Chile in 1982. During a financial crisis in Asia, observers noted urban-to-rural return migration in Thailand and Indonesia. These different adjustment processes reflect either the different institutional settings or the different nature of economic shocks that the regions faced. In any case, it is difficult to empirically identify the sources or mechanisms that significantly explain the changes of earnings and sectoral employment distribution.

Hence, for the purposes of policy intervention and impact evaluation regarding labor markets, we need to have some framework that coherently explains these changes after taking into account different institutional settings or the variable nature of economic shocks. In this paper, we attempt to construct a model to characterize the labor structure in developing countries and evaluate the impact of economic shocks on labor market variables.

The impact analysis critically depends on how we characterize labor markets. Empirical evidence suggests that a perfectly competitive labor market poorly reflects the complexity of labor
market arrangements in developing countries. We thus follow the literature and view segmented labor markets in developing countries through the lens of a multi-sector framework, where a "modern" or "productive" sector coexists with a "traditional" low-paying urban one, and a subsistence agriculture. Most of the working population is employed in the last two sectors as a result of choice, exclusion or both.¹

Few papers have attempted to model the non-competitive, multi-sector labor markets of developing countries. The first type of model to capture some of these features dates back to the dual labor market of Harris and Todaro (1970), in which a modern urban sector coexists with a traditional rural sector. In the modern sector, the wage is set above the market clearing level, thus creating unemployment. Because migration from a rural area to an urban one is costly, rural and urban wages are not equal, and migration occurs until expected returns, less the migration cost, are equalized in each area.

Fields (1996) extends the Harris-Todaro dual labor market framework to incorporate an informal or low productivity urban sector. In his model, wages set above market clearing levels create a pool of workers that queue for a formal sector job, and they can either engage in informal self-employment or remain unemployed while queuing for formal work.

Satchi and Temple (2009) characterize labor markets following the Harris-Todaro tradition of the dual labor market, where urban and rural migration take place and workers in an informal sector queue for formal sector jobs. There is no unemployment, and informality is a result of labor search frictions. Wages can be set either through bargaining or by firms using efficiency considerations.

¹For the characteristics of the informal sector, Fields (1975), Maloney (1999, 2004) and Schneider et al. (2000) are good references.
Albrecht, Navarro and Vroman (2006) (henceforth, ANV 2006), extend the studies of Mortensen and Pissarides (1994) and Pissarides (2000) by allowing for the heterogeneous productivity of workers. Workers choose a sector in which to work by comparing incomes from self-employment in the informal sector and wage work in the formal sector. The result is a market that is endogenously segmented by skills. In contrast with the aforementioned papers, ANV deviate from the dual labor market structure. Their specification of the informal sector is motivated by the fact that there is no conclusive empirical evidence to support the dual labor market formulation in developing countries. Additionally, Maloney (1999, 2004), Bosch et al. (2008) and Pratap et al. (2006) empirically show that the informal sector is not merely an undesirable queuing sector.

The models mentioned above differ not only in the number of segments, but also in the nature of segmentation. In Fields (1975), the segmentation is a result of both migration costs and a non-market clearing wage prevailing in the urban formal sector. In Satchi and Temple (2009), segmentation arises as a result of labor search frictions, and in ANV (2009), the process involves heterogeneous workers coupled with search frictions and sector-specific production technologies that endogenously segment labor markets. All of these features are present to some degree in the labor markets of developing countries, and we construct our model in such a way as to incorporate many of these features.

Our model extends ANV’s (2009) multi-sector framework by introducing the agricultural sector of Satchi and Temple (2009). First, there are many empirical evidence to support labor market segmentation in developing countries, and recent empirical findings show that informal sector jobs are preferably chosen by workers rather than are taken, as a last resort, by the workers\(^2\), who fail

\(^2\)See Maloney (1999, 2004), Bosch et al. (2008) and Pratap et al. (2006)
to search for a formal sector job but is not able to bear a high cost of unemployment. Second, Migration, especially, from rural to urban areas is a feature observed in developing countries. Economic development tends to go along with urbanization, which provides workers with more job opportunities in urban areas. Thus, the urbanization, in turn, attracts workers from rural to urban areas. We think that our model by incorporating these two features better characterize labor market in developing countries. Additionally, we calibrate our model to Nicaraguan labor market and numerically, evaluate how the labor market adjust itself when economic shock hits the economy.

Specifically, we calibrate it to the labor market in Nicaragua in 2001 and simulate economic shocks to the benchmark calibrated economy. We consider two types of shocks: one is a sector-wide productivity shock, which affects all firms of a given sector, and the other type is an idiosyncratic shock, which increases the rate of job destruction in given sectors. Our main findings are as follows. First, workers with similar characteristics can be found in different sectors with different earnings. Second, sectors in our model are segmented by skill, so that more productive workers are found in the urban formal sector while workers in the informal and agricultural sector are relatively less productive. Third, our model is capable of generating labor relocation, and the direction of the relocation is dictated by both originating sector and the type of economic shock. For example, when a negative formal sector TFP shock hits the model economy, some workers who only accepted formal sector jobs before the shock, are now willing to take informal sector jobs rather than remain unemployed.

Our main contribution is that we substantially extend the model of ANV (2009) by adding an agricultural sector and describing urban-to-rural return migration. As we argued above, rural-to-
urban migration is well observed along the economic development in the developing countries. Additionally, urban-to-rural return migration is an important labor market adjustment mechanism when economic shock hits labor market, and the return migration was empirically observed in Thailand and Indonesia during the Asian financial crisis. Finally, we contribute to the literature by calibrating our model and numerically describing underlying mechanism that works on the labor market adjustment process when economic shock hits the economy.

In the next section, we construct our model. We discuss parameterization and benchmark calibration results in the third section. Our simulation outcomes are explained in the fourth section, and we offer conclusions in the last section.

1.2 Model

Workers in agriculture receive the average product of labor, and they can migrate to urban areas for employment opportunities at a cost. When they arrive in the urban areas, they start off unemployed. The urban sector is composed of a formal wage sector and an informal one. The former is highly productive and offers bargained wages, while the latter is relatively unproductive and gives a fixed income. The model economy can face two types of shocks: idiosyncratic shocks, which hit formal or informal jobs, and sector-wide ones, which affect all jobs in the given sectors. The following sections describe our modeling strategy in greater detail.
1.2.1 Environment

Production takes place in three different sectors, the formal or high (average) productivity sector, the informal or low (average) productivity sector, and in agriculture. Agricultural output is produced in rural areas, while formal and informal jobs are located only in urban areas. Urban and rural areas are separated by distance, which make migration between rural and urban sectors costly.

There is a mass of workers $L$ who can be described by one of four conditions: i) unemployed workers looking for jobs in urban areas, ii) employed workers in the urban informal sector, iii) employed workers in the urban formal sector and iv) employed workers in the agricultural (rural) sector. We normalize the mass $L$ to unity without loss of generality.

Each firm in the urban formal sector is one job. Thus, it employs only one worker. Workers may differ in their maximum productivity. In particular, workers are identified according to a distribution function $F(y)$, $0 \leq y \leq 1$. The parameter $y$ can be loosely interpreted as the level of human capital or skill for each worker, which can be a combination of education, experience and other unobserved talents. A worker’s output in the formal sector job depends on his or her skills and on sector-wide productivity. The formal sector output production of a worker type $y$ is specified by $A_1y$, where $A_1$ is an exogenous sector-wide technology parameter.

Formal sector jobs are started at the maximum worker productivity $y$ for a worker type $y$. However, job specific shocks arrive at a Poisson rate $\lambda$ and affect the productivity of workers in that particular job. These shocks can be caused by structural shifts in demand that change the relative price of the good produced or by changes in the unit cost of production. They are real shocks associated with shifts in technology or tastes that affect that particular firm.
As in ANV (2009), we assume that productivity shocks have a common distribution $G(.)$. Once the job specific shock arrives, each worker type $y$ randomly draws a new productivity within the interval $[0, y]$. As a result, the density of productivity shocks for a worker type $y$ is $g(y')/G(y)$ for $0 \leq y' \leq y$. This means that a low productivity worker cannot turn into a high productivity worker and that the most he or she can produce is $A_1y$.

As for the productivity shocks to formal sector employment, there are two possibilities to consider. First, if the realized value of the shock of $y'$ is sufficiently low, it is mutually beneficial for both the firm and the worker to dissove their relationship. As in the standard Mortensen-Pissarides (MP) model, there is a reservation productivity $R(y)$, which depends on worker type $y$, and below which it is not worth keeping his or her job. Here, with a probability $G(R(y))/G(y)$, the idiosyncratic shock can terminate the job. Second, if the realized shock is $R(y) \leq y' \leq y$, the productivity of the worker changes to $y'$. With the probability $1 - G(R(y))/G(y)$, the job continues at the productivity level of $y'$, and the output of the worker type $y$ changes to $A_1y'$.

In the informal sector, all workers receive income $y_0$. Informal sector jobs are also subject to sectoral shocks. When the shock arrives, informal jobs are destroyed, and workers are thrown into unemployment. Shocks to informal sector jobs arrive at a Poisson rate $\delta$. Workers do not search for formal employment opportunities while they are engaged in informal production.

Workers in agriculture receive the average product of labor, and their labor income is equal to $y_a = A_a l_a^{\gamma-1}$ with $0 < \gamma < 1$, $A_a > 0$, and $l_a$ being the fraction of workers employed in agriculture. This production function simply reflects the decreasing marginal productivity of labor with a fixed amount of land. Agricultural workers do not search for jobs in urban areas when they are employed.
in agriculture. To search for an urban job, workers need to migrate to urban areas. Migration is costly, and this cost is denoted by $M$.

Unemployment is a residual state. Anyone not employed in the formal, informal or agricultural sector is defined to be unemployed and looking for a job. Unemployed workers receive a flow income of $b$, which can be equivalent to the value of leisure, unemployment benefits or transfers from other family members and friends. We assume that $y_0 > b$.

We assume that it takes time for workers to find a job and for firms to find suitable candidates to fill their vacancies. Unemployed workers find informal sector opportunities at an exogenous Poisson rate $\alpha$. When any informal opportunity arrives, they can take it or reject it.

Formal sector job opportunities arrive at the rate of $m(\theta)$, where $\theta$ characterizes market tightness, which is defined by the ratio of vacancies to unemployment. As in the standard MP model, a higher number of vacancies relative to job seekers leads to less difficulty for the unemployed to find a job and more difficult for firms to fill vacancies. As a result, the rate of arrival of formal job opportunities to unemployed workers depends on both vacancies and unemployment. We take the standard CRS matching function as in Pissarides (2000).

When workers and firms meet, the parties form a match whenever it is mutually beneficial. Both parties bargain over a wage $w(y, y)$, with the second term in parentheses reflecting worker type and with the first term standing for the current productivity of the job. After a firm and a worker are engaged in output production at a bargained wage $w(y, y)$, the wage is renegotiated whenever a shock arrives and the match remains worth keeping.

\footnote{On the other hand, fewer vacancies and more unemployed workers imply that it is harder to find jobs and easier to fill in vacancies.}
In addition to idiosyncratic shocks, there are unanticipated sector-wide shocks. These shocks are embodied in the change in the productivity levels (i.e., $A_1, A_α$, and $y_0$). They can be interpreted as changes in tastes or technology that affect the whole economy, rather than a particular firm (job). Figure 1 shows the structure of the labor market and worker flows across sectors.

### 1.2.2 Workers

For any given values of $A_1, A_α$, and $y_0$, employment composition across sectors and regions reflects workers’ optimizing behaviors.

For a worker type $y$, the value of being unemployed $U(y)$ is given by

$$rU(y) = b + α \max[N_0(y) - U(y), 0] + m(\theta) \max[N_1(y, y) - U(y), 0]$$

(1.1)

The discounted value of unemployment depends on the flow income $b$ plus the option values of being employed. At rate $α$, an unemployed worker meets an informal sector opportunity. When an informal opportunity arrives, she takes the informal job opportunity only if the surplus value (i.e., $N_0(y) - U(y)$) is positive. Similarly, formal job matching is made at the rate of $m(\theta)$.

The value of an informal sector job $N_0(y)$ is characterized by

$$rN_0(y) = y_0 + δ(U(y) - N_0(y))$$

(1.2)

Workers in the informal sector receive the flow income $y_0$ regardless of their level of skill. At rate $δ$, an informal sector job is destroyed, and the worker moves to the unemployment state.

The value of working in agriculture $N_α(y)$ is given by
\[ rN_a = y_a = A a_{\alpha - 1}. \] (1.3)

The flow value \( rN_a \) is the flow income from agricultural activities, as we assume that no idiosyncratic shocks affect agricultural employment.

Finally, the value of a formal sector job \( N_1(y', y) \) is given by

\[
rN_1(y', y) = w(y', y) + \lambda \frac{G(R_w(y))}{G(y)} (U(y) - N_1(y', y)) + \\
\lambda \int_{R_w(y)}^{y} (N_1(x,y) - N_1(y', y)) \frac{g(x)}{G(y)} dx. \] (1.4)

A worker of type \( y \) with current productivity \( y' \) receives a flow income equal to her wage \( w(y', y) \). The worker can lose her job when an idiosyncratic shock arrives with the rate of \( \lambda \) and a random draw of new productivity \( x \) falls below \( R_w(y) \). In this case, she has a surplus of \( U(y) - N_1(y', y) \). On the other hand, she can keep the job even with the arrival of the shock when the new productivity \( x \) is greater than or equal to \( R_w(y) \). The expected net gain in this case is

\[
\int_{R_w(y)}^{y} [N_1(x,y) - N_1(y', y)] \frac{g(x)}{G(y)} dx.
\]

### 1.2.3 Formal Sector Firms

Firms post vacancies if it is profitable to do so. When they meet a potential worker, they decide whether or not they consummate the match with her or search for another worker. When the idiosyncratic shocks hit formal jobs, firms have to decide whether they terminate matches with their workers or continue at the new productivity levels.
The value of a job $J(y', y)$ to a firm is the present discounted value of expected profit. As long as this value is positive, firms maintain matches with their workers. As $J(y', y)$ is increasing in $y$, there is a value of productivity $x = R(y)$, below which it is better to dissolve the match. $J(y', y)$ is given by

$$rJ(y', y) = A_1y' - w(y', y) + \lambda \frac{G(R(y))}{G(y)}(V - J(y', y)) + \lambda \int_{R(y)}^{y} (J(x, y) - J(y', y)) \frac{g(x)}{G(y)} dx.$$  \hspace{1cm} (1.5)$$

The discounted value of the job is the flow profit $A_1y' - w(y', y)$ plus expected surplus values. As in the case of formal workers, the surplus is determined by a shock process. When an idiosyncratic shock $\lambda$ arrives, firms keep the matches with their workers as long as random draws of new productivity levels $x$ are greater than or equal to the reservation productivity $R(y)$.

The value of holding a vacancy is given by

$$rV = -c + \frac{m(\theta)}{\theta} E \max[J(y, y) - V, 0].$$  \hspace{1cm} (1.6)$$

Firms pay the flow cost $c$ to keep their vacancies open. At the rate of $m(\theta)/\theta$, formal firms meet a worker. As long as the expected surplus from a match $J(y, y) - V$ is positive, firms form matches with their potential workers. Otherwise, firms keep vacancies open with no surplus. The expected value of $J(y, y) - V$ depends on the rate at which firms meet a worker of type $y$, which is determined by the distribution of $y$ in the unemployment pool. We discuss this expected value in more detail when we describe our calibration strategy.
In equilibrium, firms post vacancies until it is no longer profitable to do so. Since we assume free entry in equilibrium, the value of vacancy has to be zero.

### 1.2.4 Wage Bargaining

We assume that wages are negotiated between firms and workers in the formal sector. We follow the standard literature and assume that the wage is determined through generalized Nash bargaining. Nash bargaining guarantees that the outcome is mutually optimal in sharing total surplus from a match.

The match surplus arises because, on the one hand, workers cannot find jobs instantly, and on the other hand, it takes time for firms to fill the vacancies. Therefore, there is an outside option value from a match failure. The difference between the option value and the value from a successful match generates the surplus. A solution to the Nash bargaining problem is the wage at which the surplus from a match is optimally split between a firm and a worker, according to the given bargaining power of each party.

\[
\max_{w(y',y)} \left[ N_1(y',y) - U(y) \right]^{\beta} \left[ J(y',y) - V \right]^{1-\beta}. \tag{1.7}
\]

Since \( V = 0 \) in equilibrium, we can easily verify that the solution to the problem above is

\[
w(y',y) = \beta A_1 y' + (1 - \beta) r U(y).
\]

Simply put, the wage is a weighted average of the total formal output and the outside option value of unemployment.
1.2.5 Reservation Productivities

The Nash bargaining problem has an important property. Workers and firms under wage contracts always agree on the level of productivity below which it is mutually agreeable to revoke these contracts. In other words, firms and workers of type \( y \) have the same reservation productivity (i.e., \( R_w(y) = R(y) \)), such that the total surplus in (1.7) at \( y' = R(y) \) is equal to zero. This means that the reservation productivity \( R(y) \) can be obtained by setting the surplus of either the worker or the firm to zero.\(^4\)

Setting \( J(R(y), y) = 0 \), we have

\[
R(y) = \frac{(r + \lambda)G(y)rU(y) - \lambda A_1 \left( \int_{R(y)}^{y} (1 - G(x))dx - (1 - G(y))y \right)}{[rG(y) + \lambda]A_1}. \tag{1.8}
\]

For any given worker type \( y \), the left hand side is increasing in \( R(y) \), whereas the right hand side is decreasing. Thus, a unique solution exists. Note that the reservation productivity decreases in the formal sector productivity \( A_1 \), so that economic growth, captured by an increase in this technology parameter, has more workers and firms engaged in production.

1.2.6 Cutoff Productivities and Unemployment Values

Combining (1.2) and (1.4) with (1.1), we find an expression for \( rU(y) \) in terms of the reservation productivity and parameters, and \( rU(y) \) can be shown to be continuous and increasing in \( y \).

\(^4\)This holds true as long as wages are negotiated. When we introduce a wage floor such as a minimum wage, we see that this condition no longer holds for all worker types.
When an employment opportunity, either from the formal sector or the informal one, arrives for unemployed workers, they must decide whether or not to accept it. These are straightforward binary decisions as expressed in (1.9). Given our assumption that \( y_0 > b \), workers can be classified into three categories: (i) workers who accept only formal employment, (ii) those who accept both formal and informal offers, and (iii) those who accept only informal employment.

Formal workers with a low productivity \( y \) derive low value from their jobs. When \( y \) is sufficiently small, it is not worth being employed in the formal sector because they can expect higher income from informal sector jobs. These workers never accept a formal sector job. For other worker types, the value of being employed in the formal sector is large but is not large enough for them to stay unemployed and wait for a formal sector job opportunity to arrive. As a result, as soon as they find an opportunity to take an informal job, they accept it, but if a formal sector job opportunity arrives, they also take it. For workers with very high productivities, it is worth staying unemployed until formal job opportunities arrive, as their expected value from a formal job is very high compared with that from an informal sector job.

Two cutoff productivities are defined by the marginal worker type between the categories described above. High cutoff productivity \( y^{**} \) describes the type of marginal worker between the first two categories—(i) and (ii), while the low cutoff value \( y^{*} \) is the marginal worker’s type between the last two categories—(ii) and (iii).
Workers in different categories have different expressions for both reservation productivities and unemployment values. Unemployment values are given by

\[ rU(y) = k_1 + k_2[G(y)y + \frac{\lambda}{r + \lambda} \int_R^y (1 - G(x))dx]. \] (1.10)

Depending on a worker’s productivity \( y \), \( k_1 \) is given by

\[
\begin{align*}
    k_1 &= \frac{b(r+\delta)+\alpha y_0}{r+\delta+\alpha} \quad y \leq y^* \\
    &= \frac{[b(r+\delta)+\alpha y_0](rG(y)+\lambda)}{(r+\delta+\alpha)(rG(y)+\lambda)+G(y)(r+\delta)\beta} \quad y^* \leq y \leq y^{**} \\
    &= \frac{b(rG(y)+\lambda)}{\lambda + (r+\beta)\beta G(y)} \quad y \geq y^{**}.
\end{align*}
\]

Similarly, \( k_2 \) is given by

\[
\begin{align*}
    k_2 &= 0 \quad y \leq y^* \\
    &= \frac{(r+\delta)\beta A_1}{(r+\alpha)G(y)\beta} \quad y^* \leq y \leq y^{**} \\
    &= \frac{m(\beta)\beta A_1}{\lambda + (r+\beta)\beta G(y)} \quad y \geq y^{**}.
\end{align*}
\]

For \( y \leq y^* \), \( k_2 = 0 \), which implies that workers with a productivity level below \( y^* \) do not have an option value of being employed in the formal sector. The value of unemployment for \( y < y^* \) depends only on the flow income from unemployment and informal employment.

To find the low cutoff productivity, we use the fact that a worker of type \( y^* \) is indifferent between unemployment and a formal sector job, and \( y^* \) is given by

\[
y^* = \frac{b(r+\delta) + \alpha y_0}{(r+\alpha+\delta)A_1} - \frac{\lambda}{(r+\lambda)G(y^*)} \int_{R(y^*)}^{y^*} (1 - G(x))dx.
\] (1.11)

Workers with \( y = y^{**} \) are indifferent between informal sector work and unemployment, and \( y^{**} \) is given by
\[ y^{**} = \frac{(rG(y^{**}) + \lambda)}{\beta m(\theta)G(y^{**})A_1} (y_0 - b) + \frac{y_0}{A_1} - \frac{\lambda}{(r + \lambda)G(y^{**})} \int_{R(y^{**})}^{y^{**}} [1 - G(x)] dx. \quad (1.12) \]

### 1.2.7 Steady-State Conditions in Urban Areas

The model’s steady-state conditions allow us to solve for the distribution of workers of type \( y \) between sectors. Let \( u(y) \), \( n_0(y) \), and \( n_1(y) \) be the fractions of workers of type \( y \) in unemployment, informal sector employment, and formal sector employment, respectively, with \( u(y) + n_0(y) + n_1(y) = 1 \). We implicitly assume for now that the fractions of workers of type \( y \) in agricultural employment is zero whenever \( u(y) + n_0(y) + n_1(y) = 1 \). This implicit assumption will be straightforward after no-migration conditions are explained in the following sections.

Workers of type \( y < y^* \) flow back and forth only between unemployment and employment in the informal sector. The steady-state condition for these workers is given by

\[ \alpha u(y) = \delta (1 - u(y)). \]

We then have

\[
\begin{align*}
    u(y) &= \frac{\delta}{\delta + \alpha} \\
n_0(y) &= \frac{\alpha}{\delta + \alpha} \\
n_1(y) &= 0.
\end{align*}
\quad (1.13)
\]

There are two steady-state conditions for workers with \( y^* \leq y \leq y^{**} \); (i) the flow between unemployment and informal sector employment and (ii) the flow between unemployment and formal
sector employment.

\[ \alpha u(y) = \delta n_0(y) \quad (1.14) \]

\[ m(\theta)u(y) = \lambda \frac{G(R(y))}{G(y)} (1 - u(y) - n_0(y)) . \]

Combining these conditions, we have

\[ u(y) = \frac{\delta \lambda G(R(y))}{\lambda (\delta + \alpha) G(R(y)) + \delta m(\theta) G(y)} \]

\[ n_0(y) = \frac{\alpha \lambda G(R(y))}{\lambda (\delta + \alpha) G(R(y)) + \delta m(\theta) G(y)} \quad (1.15) \]

\[ n_1(y) = \frac{\delta m(\theta) G(y)}{\lambda (\delta + \alpha) G(R(y)) + \delta m(\theta) G(y)} . \]

Finally, workers with \( y > y^{**} \) have one steady-state condition that describes the flow between unemployment and formal sector employment.

\[ m(\theta)u(y) = (1 - u(y)) \lambda \frac{G(R(y))}{G(y)} . \]

This implies

\[ u(y) = \frac{\lambda G(R(y))}{\lambda G(R(y)) + m(\theta) G(y)} \]

\[ n_0(y) = 0 \quad (1.16) \]

\[ n_1(y) = \frac{m(\theta) G(y)}{\lambda G(R(y)) + m(\theta) G(y)} . \]

Total unemployment is obtained by aggregating across the urban population, as there is no unemployment in agriculture in this construction.

\[ u = \int_{y_r}^{1} u(y) f(y) dy . \quad (1.17) \]
Analogously, we can compute the total share of workers in the informal and formal sectors. The share of workers in the formal sector is given by

\[ n_1 = \int_{y_r}^{1} n_1(y) f(y) dy, \]

and the share of workers in the informal sector is

\[ n_0 = \int_{y_r}^{1} n_0(y) f(y) dy. \]

By this construction, \( n_0 + n_1 + u + F(y_r) = 1. \)

### 1.2.8 No-Migration Condition(s)

We assume that if workers migrate to urban areas, they start off unemployed and that workers in the urban sectors consider migrating back to rural areas only when they find themselves unemployed. If urban workers migrate to rural areas, they can start working as agricultural workers immediately. Thus, agricultural work is always available. This means that the migration decision is essentially one of comparing the value of being unemployed \( U(y_r) \) with that of being in the agricultural sector \( N_a(y_r) \) and the migration cost \( M \). In equilibrium, no migration takes places, which means that for all workers located in rural areas,

\[ rN_a(y_r) \geq rU(y_r; \theta) - M \quad (1.18) \]

and for all workers located in the urban areas

\[ rU(y_r; \theta) \geq rN_a(y_r) - M. \quad (1.19) \]
Unilateral (Rural-to-Urban) Migration

With numerical simulations in mind, we think of the benchmark as the state before recessionary economic shocks. Thus, urban sectors typically attract workers from rural areas in the process of economic development. Then, the no-migration condition requires

\[ rN_a(y_r) = rU(y_r; \theta) - M, \]  

(1.20)

and the share of workers in agriculture is \( l_a = F(y_r) \), which, combined with the fact that \( rN_a(y_r) = A_a l_a^{\gamma - 1} \), implies

\[ A_a [F(y_r)]^{\gamma - 1} = rU(y_r; \theta) - M \]  

(1.21)

Depending on parameter values, specifically on whether \( y_r \) is greater or less than \( y^* \), \( U(y_r) \) may or may not depend on \( y_r \) and \( \theta \). If \( y_r < y^* \), then \( U(y_r) \) is a constant equal to \( \frac{b(r + \delta) + \alpha y_0}{r + \delta + \alpha} \) for all \( y \in [y_r, y^*] \), but if \( y_r \geq y^* \), then \( U(y_r) \) is increasing in both \( y_r \) and in \( \theta \). In this latter case, equation (1.21) defines a locus of \((y_r, \theta)\) combinations that are consistent with the no-migration condition. If \( y > y^{**} \), there would be no informal sector, and we exclude this possibility from our analysis. Figures 2 and 3 illustrate both cases.

The Job Creation Condition

We use the free-entry condition to close the model and determine equilibrium labor market tightness. Setting \( V = 0 \) gives

\[ c = \frac{m(\theta)}{\theta} E \max[J(y, y), 0]. \]
To determine the expected value of meeting a worker, we need to account for the fact that the density of types among unemployed workers is contaminated. Let \( f_u(y) \) denote the density of types among the unemployed. Using Bayes’ Law, we have

\[
f_u(y) = \frac{u(y) f(y)}{u}.
\]

The free-entry condition can thus be rewritten as

\[
c = \frac{m(\theta)}{\theta} \int_{\max[y_r, y^*]}^{1} J(y, y) \frac{u(y)}{u} f(y) dy.
\]

This expression takes into account that no jobs will be created for \( y < y^* \), so that the lower limit of integration is the highest between \( y_r \) and \( y^* \). If \( y_r < y^* \), then the lowest level of skills for which it is worth engaging in production is \( y^* \). If \( y_r > y^* \), then the lowest worker type in the urban sector is \( y_r \). After substitution for \( J(y', y) \) evaluated at \( y' = y \), this becomes

\[
c = \frac{m(\theta)}{\theta} \left( 1 - \beta \right) A_1 \int_{\max[y_r, y^*]}^{1} \left[ y - R(y) \right] \frac{u(y)}{u} f(y) dy.
\]

Equation (1.22) is equivalent to the job creation condition in Mortensen and Pissarides (1993), and it is an equation of \( \theta, R(y) \) and \( y_r \). It only makes sense if the right-hand side is positive. Because \( J(y^*, y^*) = 0 \) and \( J(y, y) \) is increasing in \( y \) for \( y \geq y^* \), a necessary condition for equation (1.22) to have a solution is \( \max\{y^*, y_r\} < 1 \).

**Equilibrium with Rural-to-Urban Migration Only**

The job creation condition and the no-migration condition jointly determine the equilibrium levels of \( \theta \) and \( y_r \). There are two cases depending on whether \( y^* > y_r \) or \( y^* < y_r \). Figure 4 and 5 illustrate...
each case. Our equilibrium concept is more general than the one in ANV (2009) in that they only consider the case \( y^* > y_r \). Thus, they think that there always exist sufficiently low-skilled workers, who are engaged only in the informal sector. By incorporating the case \( y^* < y_r \), we include urban labor markets where all unemployed urban workers are potential candidates for formal sector jobs. All the other components of equilibrium are the same as in ANV (2009).

1.2.9 Bilateral Migration

As observed in Thailand and Indonesia during the Asian financial crisis, return migration may occur in the labor market adjustment process against economic shocks. Thus, we need to specify no-migration conditions in which workers are allowed to migrate in both directions. Although economic shocks can relocate workers across sectors, not all shocks generate migration. Movements between urban and rural areas may take place only when the gain from migration is big enough to cover the migration cost.

Thus, equilibrium in this case requires that for all workers located in the rural areas,

\[
r_{N_a}(y_r; y^p_r) \geq r_U(y_r; y^p_r, \theta) - M
\]  

(1.23)

and for all worker types located in the urban areas,

\[
r_U(y_r; y^p_r, \theta) \geq r_{N_a}(y_r; y^p_r) - M.
\]  

(1.24)

Notice that the no-migration condition is characterized as conditional on the threshold value \( y^p_r \). This \( y^p_r \) is needed to pin down \( y_r \), and we may consider \( y^p_r \) as the threshold value of the no-migration
condition(s) before an economic shock.

Let \( y^a \) be the value of \( y \) for which equation (1.23) holds with equality

\[
r_{N_a}(y^a) = r_U(y^a) - M
\]  

(1.25)

and let \( y^c \) be the value of \( y \) for which for which equation (1.24) holds with equality

\[
r_{N_a}(y^c) = r_U(y^c) + M
\]  

(1.26)

Figure 6 illustrates the above equations for the case where \( y^a < y^* \).

Two things are worth noting. For any equilibrium, i) \( y^c < y^a \), and ii) there are no workers with \( y < y^c \) located in the urban areas and with \( y > y^a \) located in rural areas. However, workers with \( y^c < y < y^a \) may be located in either rural or urban areas and still behave optimally in their choice of location.

Once a shock hits the economy, \( y_r \) is necessarily \( y^c \leq y_r \leq y^a \) in a new equilibrium, and this level of skills may imply some migration if \( y_r \neq y^p_r \). No migration occurs if \( y^c \leq y^p_r \leq y^a \) because

\[
r_{N_a}(y^p_r; y^p_r) \geq r_U(y^p_r; y^p_r, \theta) - M \quad \text{and} \quad r_U(y^p_r; y^p_r, \theta) \geq r_{N_a}(y^p_r; y^p_r) - M.
\]

In this case, \( y_r = y^p_r \), and the case is illustrated by a point like B in figure 6, where B represents the initial threshold level of skills \( y^p_r \). If, on the other hand, \( y^p_r < y^c \), then it is optimal for some workers in the urban sector to move back to agriculture\(^5\) and \( y_r = y^c \). This case is illustrated by a point like A in the figure.

Finally, if \( y^p_r > y^a \), then it is optimal for some agricultural workers to migrate to the urban areas\(^6\)

\(^5\)In this case, \( r_U(y^p_r; y^p_r, \theta) < r_{N_a}(y^p_r; y^p_r) - M \). Thus, some urban unemployed workers can be better off when they migrate into agriculture.

\(^6\)In this case, \( r_{N_a}(y^p_r; y^p_r) < r_U(y^p_r; y^p_r, \theta) - M \)
and \( y_r = y^a \).

Thus, for an initial allocation of workers determined by \( y_r^p \), the equilibrium threshold level of skill is given by:

\[
y_r = \begin{cases} 
  y^c, & y_r^p < y^c < y^a \\
  y_r^p, & y^c < y_r^p < y^a \\
  y^a, & y^c < y^a < y_r^p
\end{cases}
\]  
\( (1.27) \)

**Equilibrium with Bilateral Migration**

A *steady-state multi-sector labor search and matching equilibrium* is characterized by a labor market tightness \( \theta \) and no-migration cutoff value \( y_r \) together with a reservation productivity function \( R(y) \), unemployment rates \( u(y) \), and cutoff values \( y^*, y^{**}, y^a \) and \( y^c \) and a given latest no-migration cutoff value \( y_r^p \) such that

1. the value of maintaining a vacancy is zero.
2. matches are consummated or dissolved if and only if it is mutually profitable to do so.
3. the steady state conditions hold.
4. formal sector matches are not worthwhile for workers with \( y < y^* \).
5. informal sector matches are not worthwhile for workers with \( y > y^{**} \).
6. Given \( y_r^p \), the no-migration condition holds and is determined by \( (1.23)-(1.27) \).
1.3 Simulating the Impact of Economic Shocks

Economic shock can affect either entire economy or selected firms. For example, a financial shock can affect entire economy because generally, tightened credit markets make it more expensive to finance firms production schedules due to increase in the rental price of capital. However, the shock can negatively affect mainly exporting firms if the shock is originated in foreign importing countries and domestic economy is relatively immune to the shock. Another example may be food price shock that we observed in 2007-2008, and natural catastrophes like Haiti’s earthquake in 2010 and Tsunami triggered by Indian ocean earthquake of Sumatra in 2004. In all cases, wages and employment are adjusted to the shocks, and our purpose of simulation exercises is to quantitatively evaluate how wages and labor composition are changed when economic shock hits the economy.

If the shock affects entire economy, we view the shock as a structural perturbation on $A_1$, $y_0$, and/or $A_0$ in our model. That is, as TFP in the formal, informal and/or agricultural sector falls, all firms are either directly or indirectly affected. The economic shock can affect some firms but not all of them. In this case, job turbulence in the economy increases, which can be interpreted as an increase in the arrival rates of idiosyncratic shocks ($\lambda$ and/or $\delta$) to the formal and/or informal sector in our model.

Do the impacts of increased turbulence or reduced profitability in the formal sector trickle down to the informal sector or even to the agricultural one? Although impacts on the other sectors are quite complicated to estimate empirically, post-crisis labor statistics in developing countries show significant changes in the informal sector (i.e., an increase in informal employment share, reduced employment duration, etc.). Because we aim to measure direct and indirect trickle-down effects of
the economic shock in labor markets, we evaluate the effect by simulation in our model.

1.3.1 Data

We use several data sources to construct target statistics for our calibration. The main source is from Nicaragua’s 2001 annual Household Survey (Encuesta de medicion del Nivel de Vida (EMNV)), in which we find the share of workers in each sector and their level of income and education. We also use data from the Food and Agriculture Organization (FAOSTAT) to calibrate the agricultural production function.

Before we discuss the parameterization and calibration strategy of our model, we need to define which employed worker types are classified into each sector and how broad a concept of unemployment we should employ. We assign family enterprise workers and individually self-employed and non-regulated wage workers\(^7\) outside of agriculture to the informal sector. Regulated wage workers are assigned to the formal sector, and agricultural workers other than employers are all assigned to the agricultural sector. Employers in both agriculture and non-agriculture sectors are left out all together. We think that they are better understood as firms rather than as workers. Overall, informality is distinguished from formality in terms of the degree of labor productivity, on the one hand, and the compliance with labor regulations, on the other.

The transition rate from non-employment to employment is similar among the discouraged, temporarily inactive and (narrowly defined) unemployed, and so is the rate of transition in the opposite direction. This suggests that those unemployed workers may not be that different from

\(^7\)Non-regulated wage workers are those who are not subject to social security payment, while the regulated wage workers are the opposite.
the discouraged and temporarily inactive ones in terms of the willingness to work. We thus use a broad concept of unemployment that encompasses discouraged and temporarily inactive workers. While the narrow unemployment rate is around 4%, the broad unemployment rate is 12%. Table 1 shows the share of workers in each labor market state.

1.3.2 Parameterization

We opt for normalizing some of the parameters to reduce computational burdens. The leisure value of unemployment is normalized to zero, or the parameter $b$ is fixed at zero; the bargaining power of formal sector workers is set at the standard value $1/2$, or $\beta$ is equal to $1/2$; $r$, the discount rate, is set at 0.04.

We first specify parameters of the shock process $G(y')$ and the skill distribution $F(y)$. We assume that the idiosyncratic shock process is characterized by a uniform distribution with support of $[0,y]$. In other words, when a shock arrives, the new level of productivity of a worker type $y$ lies between 0 and $y$, with equal probability of ending up anywhere in the support.

We assume that the skill distribution follows the $\beta$-distribution. There are two reasons for this choice. First, the $\beta$-distribution is defined on a finite support so that we can control for extreme values of skill, and second, the $\beta$-distribution is flexible in that it displays a broad range of shapes depending on the values of two distributional parameters $(a_\beta,b_\beta)$. We proxy the skill by the normalized values of education, and the first two moments of the values of education are sufficient to calculate the two distributional parameters and thereby, to define the distribution itself.

---

8 See Appendix A for the logic in greater detail
9 The level of education is normalized by the maximum level in the data.
Earnings in agriculture are given by the average product of labor, which is rewritten below:

\[ y_a = A_a l_a^{\gamma-1}. \] (1.28)

We estimate the elasticity of agricultural labor (\( \gamma \)) from longitudinal data on agricultural productivity and employment from 1961 to 2005.\(^{10}\) We construct the series of the working age population in the agricultural sector from the FAOSTAT. We estimate the following simple model,

\[ \log Y_a = c + \gamma \log L_a, \]

and our estimate of \( \gamma \) is 0.628.

Using the fact that the income from agriculture is given by (1.28), we back out \( A_a \) in a way that is consistent with our data for year 2001. The household survey EMNV enables us to calculate the average labor income in agriculture \( (y_a) \) and the share of the workers in agriculture \( l_a \). The average income in agriculture is C$10,717 córdobas, and \( l_a \) is 27.10% as shown in Table 1. This allows us to back up \( A_a \). Finally, when we know \( \gamma, A_a, \) and the average agricultural income, \( l_a = 27.1\% \) in Table 1, this implies that \( y_r \) is determined at \( l_a = F(y^r) \).

The informal labor market is specified by the three parameters \( \{\alpha, y_0, \delta\} \). The average income in the informal sector \( (y_0) \) is estimated from the EMNV and is \( y_0 = C$9,767 \) córdobas. Because the rate of the informal job destruction (\( \delta \)) is assumed to follow the Poisson distribution, the inverse of the rate (i.e., \( 1/\delta \)) is equivalent to the average duration of holding an informal sector job. Since we can calculate the duration of informal sector jobs from the EMNV, we can back out \( \delta \).

\(^{10}\)We use the longitudinal series of the agricultural output \( Y_a \) published by the Central Bank of Nicaragua.
We still have five more parameters to determine.\textsuperscript{11} As far as we know, we do not have relevant data available to fix these parameters. We thus calibrate our model to choose the values of the parameters in such a way as to best match the following target statistics.

We choose the share of workers in formal and informal sector jobs ($N_0$ and $N_1$ respectively), the average wage $\bar{w}$ in the formal sector, and the average duration of formal sector jobs $d_1$, for which our model gives explicit expressions\textsuperscript{12}, and we choose values observed in the EMNV. These targets are chosen because our model can describe labor markets along those dimensions; examples are labor market composition, sectoral employment duration and sectoral labor income distribution. The target statistics from the EMNV are listed in Table 2.

Because the model is able to generate those targets without knowing the migration cost $M$ and the cost of vacancy $c$, we first calibrate the model without the job creation and no-migration conditions to find the values of $\alpha, \lambda, A_1$, and $m(\theta)$. Once these parameters are determined within the model, the free-entry condition together with the matching specification determines the parameter $c$, while $M$ is determined by the no-migration conditions.

\subsection*{1.3.3 Results}

Table 3 shows the resulting parameter values, and Table 4 shows the values of endogenous variables. There is a huge wage difference ($\bar{w}$ vs $y_a$) between the urban formal and the rural agricultural

\textsuperscript{11}The list is as follows: the arrival rate of the informal employment opportunities ($\alpha$); the arrival rate of the idiosyncratic shock ($\lambda$) in the formal jobs; the formal sector productivity ($A_1$); the migration cost $M$; and the cost of posting a vacancy $c$.

\textsuperscript{12}See Appendix B.
workers, as there is a difference in the level of technology between the rural agriculture sector \( (A_a) \) and the urban formal sector \( (A_1) \). As a result, the sector in which one works can be an important indicator of one’s earning capacity (or individual skill in our model). Given the idiosyncratic shock arrival rate \( \lambda = 0.2342 \), formal sector jobs are expected to face a shock every four years. Some might be destroyed and others not, depending on the value of a new productivity draw. However, we are not able to observe this level, due to the lack of related information from industrial statistics. The rate of shock arrival to informal sector jobs is around half of the value of \( \lambda \). This does not mean that informal jobs are more secure, as they are destroyed whenever a shock arrives. In fact, formal sector jobs are more stable in the sense that the duration of formal sector employment is, on average, longer than the duration of informal sector employment by three and half years, which is directly verifiable from the duration data we used to construct duration targets.

Workers with more than 8 years of education search for jobs only in the formal sector, and their share of total formal sector employment is about 57%. According to the EMNV, the average education of formal sector workers is around 8 years, and about 42% of formal sector workers have more than 8 years of schooling. Our benchmark value of \( y^* \) suggests that workers with less than complete primary education (6 years) will not search for formal employment opportunities. The data suggests that only about 30% of informal sector workers have more than 6 years of education.

The no-migration condition suggests that 3 years of education represent the minimum level required to migrate into urban areas. About 55% of workers in agriculture have less than 3 years of education.\textsuperscript{13} Segmentation by skill thus seems to be a clear feature in the Nicaraguan data.

\textsuperscript{13}The percentage of workers with less than 3 years of education falls to 27% in the informal sector and 13% in the formal sector.
1.4 Simulation Results

We simulate the effects of i) 5% and 10% reductions in both the formal sector and the informal sector TFPs respectively, a 10% reduction in $A_1$ and 5% increase in $y_0$, ii) 5% and 10% increases in formal sector job turbulence $\lambda$ respectively, iii) an increase in informal job turbulence $\delta$, and iv) several combinations of the above.\(^{14}\) The results are listed in Table 5 and Table 6.

Several results are worth highlighting. First, as shown in Table 5, the fall in the formal sector productivity $A_1$ (Simulation 1) reduces the formal employment regardless of the direction of the changes in $y_0$. If the shock we consider mainly affects firms in the formal sector (especially, exporting firms), the impact to informal sector firms should be made through indirect channels, and the magnitude of negative impact should be smaller than or equal to the change in $A_1$. Thus, the first two results in Simulation 1 suggest that with the same percentage decline in the productivity of both sectors or with the urban sectoral productivity shock, the impact is severe in the formal sector, and some of formal sector workers are relocated to the informal sector.

The third exercise in Simulation 1 describes a case where the productivity in the informal sector increases. This case may be justifiable because labor productivity in the informal sector can go up when some relatively productive formal sector workers may be relocated to the informal sector from the negative TFP shocks in the formal sector. The third result in Simulation 1 suggests that the magnitude of labor relocation from the formal sector becomes larger as the value of informal sector jobs increases, in contrast with the other two cases in Simulation 1. Simultaneously, we observe

\(^{14}\)In other words, we view the financial shocks as if they are equivalent to decreases in formal sector productivity, either an increase or decrease in informal sector productivity, an increase in job turbulence or some combinations of these.
from Table 6 that per capita output\textsuperscript{15} falls, ranging from 3% to 9% in Simulation 1. Therefore, a 1% decline in total output would be associated with a reduction in formal jobs of approximately 0.3% in the first two cases and 6% in the third one.

Formal sector employment is relatively sensitive to changes in $A_1$ and $y_0$, whereas both total employment and total output are much less responsive. The reason for this is that many workers, in the face of negative shock on $A_1$, become indifferent between formal and informal sector jobs, so that workers who searched only for formal sector jobs before the shock are willing to accept informal sector jobs in the post-shock state of the economy.\textsuperscript{16} This is a direct consequence of no rigidities in wages so that wages can absorb part of the decrease in productivity in the formal sector on one hand and low income of the unemployed (i.e. $b = 0$) on the other. Consequently, changes in unemployment and total output are negligible because more workers are attracted to the informal sector.

In all cases in Simulation 1, the change in unemployment is modest relative to the labor relocation from the formal sector to the informal sector. For example, the second result in Simulation 1 suggests that 85% of the relocated workers from the formal sector are reemployed in the informal sector, whereas only 15% of them are unemployed. The level of labor relocation from the formal sector critically depends on the directional change of informal sector productivity. If both $A_1$ and $y_0$ decline, the role of the informal sector as a buffering sector becomes weaker, such that the labor relocation from the formal to the informal sector is relatively small. The second result in Simulation 1 shows a 1% expansion of the informal sector as a result of 10% productivity decline in

\textsuperscript{15}Per capita output is computed as an average income without including the cost of vacancy.

\textsuperscript{16}This change in job search behavior is from relatively large increases in $y^{**}$ in our model.
both sectors. If we reduce the formal sector productivity by 10% while holding $y_0$ at the benchmark, the informal sector expands by 4%. Furthermore, the third result in the simulation shows the expansion of the informal sector more than 10% with 10% reduction in $A_1$ and 5% increase in $y_0$.

The second simulation explores the effects of idiosyncratic individual skill shocks in the formal sector; that is, we experiment with the impacts of a shock that affects some firms in a sector but not all of them. The skill shocks measure the degree of turbulence in the formal sector. Since the idiosyncratic formal shocks affect only some individual workers, overall labor markets are much less affected. However, it is important to highlight that the calibrated turbulence parameter is rather small at the benchmark. The rate of the shock is 0.2342, so a 10% increase means that, on average, a firm faces a shock every 3.8 years rather than every 4 years.

The effect of the higher turbulence is to shed workers from the formal sector into unemployment. Consequently, unemployment rate increases together with shortened duration of formal sector jobs, as shown in Simulation 2 in Table 5 and Table 6. In this case, we also observe labor relocation from the formal to the informal sector because the value of a formal sector job declines.

The third simulation attempts to isolate the effects of the increased informal sector job turbulence $\delta$. Again, the benchmark calibrated value of the shock is rather small.\(^\text{17}\) What we emphasize in this simulation is that the turbulence in the informal sector alone can relocate workers across sectors.

A 10% increase in the informal sector turbulence $\delta$ increases the formal sector employment and total unemployment. Since the value of informal jobs declines, more workers search for for-

---

\(^{17}\)The benchmark value of $\delta$ is 0.176. In other words, the shock arrives to workers every 5.6 years on average.
mal sector jobs.\textsuperscript{18} The greater turbulence increases the unemployment rate, as informal jobs are destroyed more frequently. Finally, its impact on welfare is negative, as per capita output decreases by slightly less than 1%.

Simulations 4 and 5 allow multiple sources of shocks. The main implication is that the effect of shocks on the formal sector is overall more severe than that on the informal sector. Consequently, some formal sector workers are relocated into either informal sector or unemployment. However, the magnitude of the relocation critically depends on whether the informal sector productivity increases or decreases in the adjustment process following the shock. Overall, unemployment rates in all our simulation exercises increase as job turbulence in each sector becomes stronger. However, quantitatively, the rates increase at best 2% in our simulation exercises, as shown in Table 5.

The biggest effects on wages are obtained when the productivity in both sectors declines and the turbulence in both sectors increases. In this case, the employed workers in the formal sector become less productive and thus earn less from their jobs. This is also the case in the informal sector. Because the buffering role of the informal sector is weakened in this case, formal sector workers would rather remain in the same sector and take the reduced wages than switch to informal sector jobs.

It is important to highlight that our model does not consider congestion externalities by the increase in the number of workers willing to accept informal jobs. Nor does the model consider effects of the inflow of workers to the informal sector on informal earnings. If there are congestion externalities or the larger supply of workers to informal sector reduces earnings, a shock to the

\textsuperscript{18}In our model, both cut-off values $y^*, y^{**}$ decrease, such that more workers become potential formal sector workers.
formal sector might generate a larger relocation of workers to agriculture, as well as more unemployment, as the returns to informal employment will decrease. Thus, the economic shock might have a larger effect on backward migration.

1.5 Conclusions

This paper presents a multi-sector labor search and matching model for developing countries. Workers in our model are heterogeneous in terms of their skill and are located in one of the three job sectors or in unemployment. There is an entry barrier between urban sectors and agriculture, such that workers must pay a migration cost when they migrate into the urban area from the rural one. The governing principle behind our model is based on the fact that sectors are segmented by the level of sectoral productivities from different forms of earnings functions, while heterogeneous workers are endogenously sorted into these sectors. This sorting mechanism is such that high-skilled workers tend to be located in the relatively more productive and urban formal sector; on the other hand, low-skilled workers are found in the agricultural or urban informal sectors. However, some workers in the urban area are indifferent between formal and informal jobs, due to costly job search frictions.

We use our model to assess the potential effects of the economic shock. The first challenge is how to model the economic shocks and then to simulate the impacts of the shocks. For example, how much should the productivities in the formal and other sectors be changed? Most analyses reduce the effects of the shocks to a measure of growth lost or reduction of GDP. Although both measures are good indicators for impacts, they are themselves the results of the shocks to exoge-
nous variables such as financial instrument variables, foreign demand, etc. We thus construct the model to describe the underlying mechanism that links the shocks to its endogenous outcomes (i.e., GDP per capita, employment and earnings).

Most studies use a measure of projected GDP growth together with employment elasticity to forecast the effects on labor markets. The elasticity is a summary variable of how these two variables have responded in the past to exogenous changes. There are two drawbacks to this methodology. First, the elasticity masks important labor relocation that is likely to have important effects on welfare and poverty. Second, elasticity might change due to the new conditions in the world economy. Our model does not imply constant elasticity, and it highlights the relocation of the labor and its effects on welfare and earnings.

After constructing the model, we try to ascertain in our simulations the degree to which the economic shocks would affect firms’ profitability and job turbulence. The results suggest that despite modest changes in total employment, there are either significant labor relocations or formal sector wage-cuts, depending on whether informal sector productivity increases or not.

In Nicaragua, the adverse productivity shock in the formal sector can reduce formal sector employment. If the informal economy remains unaffected, it may provide these unemployed workers with a safety net, and unemployment may not significantly increase. However, if the informal

\[ \text{Firm profitability is well captured in our model by the parameters } A_1, A_0 \text{ and } \gamma_0, \text{ and it should not be hard to find empirical data to estimate the potential effects of changes in these parameters. Job turbulence may have an empirical counterpart in phenomena such as export sector shocks, but it is hard to identify the effects of those shocks in the data. As a result, it would be necessary to have an empirically measurable definition of turbulence and to study the impacts of it in previous crises to adequately simulate it.} \]
sector is also affected, then some workers may go back to agriculture reducing the congestion in the urban labor market, which in turn reduces the impact on the unemployment rates. Thus, unemployment increases modestly at best. The way that the economic shock affects relative returns to each sector would determine the sectoral relocation of labor and the associated losses in efficiency.
Chapter 2

Validating a Multi-sector Labor Search and Matching Model for Developing Countries

2.1 Introduction

The objective of this paper is to validate the use of a multi-sector labor search and matching model in order to explain labor market structures in developing countries. Labor markets in developing countries have distinctive features; in particular, a high-skilled productive sector often coexists with a low-skilled labor intensive sector and a subsistence agriculture. This motivates us to model labor markets in developing countries using a segmented labor market framework rather than the conventional competitive labor market treatment. Specifically, we use a model by Gutierrez, Paci and Park (2008)\(^1\), which is an extension of Albrecht, Navarro and Vroman (2009).

Our model validation will allow us to check whether the model is able to reproduce the structure

\(^1\)Hereafter, we denote them by GPP.
of actual labor markets. Specifically, we test whether the model can predict labor market outcomes not used for our model calibration. For this validation exercises, we first calibrate our model to the Indonesian labor market in 1997. We then compare our simulated outcomes to data observations. We find that the model can replicate average skill distributions across sectors, and that both sectoral employment and unemployment shares across individual skill levels fit well with the data. However, the unemployment share for highly educated workers is low in the data, whereas our model predicts a high unemployment share. Additionally, the model-generated wage distribution is bell-shaped with a high-density around the mean, but the empirical distribution is left-skewed.

Our contribution through this paper is to evaluate whether the model can potentially be used to study labor structures in developing countries. Our results show that the model is well equipped to study average behavior in segmented labor markets.

In the next section, we briefly summarize the model of GPP. In Section 3, we explain the data and target statistics for our model calibration. Benchmark parameterization will be discussed in Section 4. Our results for cross-sectional and dynamic validations will be shown in Sections 5 and 6, respectively. In Section 7, we conclude.

### 2.2 Model

Since we use the same model as in chapter 1, we will only briefly describe the model elements in this section. Our model is built on the search and matching model of Mortenssen and Pissarides (1994) and Pissarides (2000) and incorporates a multi-sector labor market. We closely follow Albrecht et al. (2009) and Satchi and Temple (2009). As discussed in Satchi and Temple (2009),
labor markets are segmented by geographic distance between urban and rural areas, and wages in the urban formal sector are determined by bargaining. We deviate, however, from the idea of dualism in the labor markets in the sense that urban informal sectors are not necessarily queuing sectors for employment opportunities in urban formal sectors.\textsuperscript{2} As discussed in GPP (2008) and ANV (2009), the urban informal sector is characterized as an optimal choice for urban workers with relatively low skills.

The agricultural sector is regarded as one big extended family, and workers are assumed to equally divide agricultural output from their labor. Agricultural workers can migrate to urban areas for employment opportunities at a cost. When they arrive in the urban areas, they start off as unemployed. Urban areas feature both formal and informal sectors. The former is highly productive and offers bargained wages, while the latter is relatively unproductive and pays a fixed income. Urban labor markets have search frictions in that it takes time for both labor-searching firms and unemployed workers to find mutually agreeable counterparts.

\section*{2.3 Data and Target Statistics}

\subsection*{2.3.1 Data Sources: Sarkernas vs. IFLS}

We use the Indonesian Family Life Surveys (IFLS) and Sakernas to construct labor market targets for our benchmark model calibration. The IFLS is a longitudinal socioeconomic survey that covers 13 out of 27 provinces and represents 83\% of the total population. The first wave was conducted in 1993 (IFLS1), and the fourth wave, fielded in 2007 through 2008, is the latest.

\textsuperscript{2} See Maloney (1999, 2004), and Pratap and Quintin (2006)
Sakernas is a main Indonesian labor force survey managed by the BPS\(^3\) to monitor and compile provincial employment statistics. It has been conducted since 1986\(^4\), and the survey covers a nationally representative sample of about 40,000 of the working-age population. However, Sakernas’ lack of self-employment income records fails to identify the informal economy of our model. Thus, we use the information about self-employment incomes from the IFLS. Although IFLS is not nationally representative, both Sakernas and IFLS show similar changes in wages and employment rates over the Asian Financial crisis.\(^5\)

However, these surveys show opposite trends in the unemployment rate. As shown in Table 7, the unemployment rate in the IFLS declined from 7.6% to 2.02%, whereas the rate had gone up from 4.70% to 6.30% in Sakernas through the crisis. As discussed in Smith et al. (2002), this discrepancy may be due to survey design. Both surveys asked a similar lead-in question. When respondents answered ‘currently not working\(^6\)’, the Sakernas stops probing further, whereas the IFLS asks two-step follow-up questions. First, respondents are asked if they had worked for pay for at least one hour. If they answer ‘No’, then the second follow-up question probes whether they had worked in a family-owned business. Therefore, the IFLS is expected to have a lower unemployment rate.

The unemployment trend in Sakernas follows the conventional wisdom that labor demand tight-

\(^3\)It is an abbreviation for *Badan Pusat Statistik* or Statistics Indonesia. This is a government agency tasked with collecting and publishing statistical data in Indonesia.

\(^4\)The survey was done quarterly before the year 1995 and annually afterwards.

\(^5\)See Smith et al. (2002)

\(^6\)Work is defined as an employed state in which one generates income or helps to generate income for at least one continuous hour.
ens when the economy is in a downturn. Consequently, more unemployed workers queue, and queue for a longer time, for employment opportunities in a recessionary state. However, we have no reasons to reject the fall in unemployment observed in the IFLS. A small decline in unemployment rate may be understood as the informal sector is expanding in economic recession in developing countries. That is, the observed increase in unpaid family business workers and in female labor force participation, particularly in the informal sector, may support the low post-crisis unemployment rate reported by IFLS.

For our target statistics, we use wage and employment statistics from the IFLS and the unemployment rate from the Sakernas. For consistency purposes, we restrict samples from the Sakernas to 17 provinces of the IFLS. As shown in Table 7, the effect of sample selection on unemployment rates is negligible.

### 2.3.2 Labor Composition and Labor Income

In Table 8, we list sectoral employment shares and the unemployment rate in the benchmark Indonesian labor market. First, we exclude employers to be consistent with the concept of workers in our model. Employers may be more like firms than workers. We also drop the *self-employed with workers*, which consists of 3.6% of the total sample. This variable is characterized in the survey at 2000 by dividing it into *self-employed with family member* and *employer with permanent workers*. We consider the former as self-employed informal workers and the latter as employers. The share of *self-employed with workers* takes only 3.6% in 1997, and *self-employed with family member* as

---

7 The reason why we use Sakernas’ unemployment rate is simply that Sakernas is nationally representative.

8 The rates in 1997 are the same while there is only 0.2% difference in year 2000.
in 2000 might have been absorbed in self-employed working alone, or unpaid family workers in 1997. Thus, we think that the informal sector captures the same characteristics of workers in both years.\footnote{In other words, if there is any discrepancy that the informal sector captures, in both years, it should be negligible.}

As discussed in GPP (2008), sample observations are classified into model variables as follows. All types of workers in the agricultural sector are assigned to agricultural employment. Informal workers are defined by self-employed working alone, self-employed with family member, unpaid family workers, or wage workers with no medical or social security contribution. Workers are employed in the formal sector if they are either government workers or wage/salary workers with social security.

Table 9 shows labor incomes across sectors. As is the general case in developing countries, an average formal sector income is highest, and agricultural income is the lowest in Indonesia.

### 2.4 Benchmark Parameterization

Since we use the same model as discussed in GPP (2008), we closely follow their parameterization strategy. In this section, we briefly explain our benchmark calibration strategies.\footnote{Refer to Appendix B in GPP (2008) for calibration methods in detail.} Our benchmark calibration is run for the Indonesian labor markets in 1997.

First, we take a discount rate of 0.1381 from the IMF data set. Workers are heterogeneous and are indexed by their skills. We assume that the skills follow a $\beta$-distribution with the support of $[0, 1]$. This distribution is estimated by taking the level of education as a proxy. Three sectors are distinguished by sector specific productivities.; $A_a$ for Agriculture, $A_1$ for the formal sector
and unity for the informal sector as a normalization. When we rewrite the agricultural production function, we have

\[ Y_A = A_a N_0^\gamma \]

We estimate \( Y_A \) and \( N_0 \) from the IFLS. We fix the agricultural labor elasticity parameter \( \gamma \) at the poor-country average as in Bravo-Ortega et al. (2004). We then determine the value of \( A_a \) within the model in such a way to match the average income in agriculture.

The model implies that workers sort themselves into sectors such that highly productive workers are more likely to be found in formal sector jobs; low productivity workers in agriculture; workers in-between are either working in the formal or informal sector, depending on their search draws. Formally, workers’ choice rules can be characterized by the following cut-off values \( y_r, y^* \) and \( y^{**} \). Agricultural workers are the ones whose skills are \( y < y_r \). Workers with \( y \in [y_r, y^*] \) choose to work only in the informal sector. If workers’ skills are greater than or equal to \( y^* \), they can work in either the formal or the informal sector, but workers with \( y \geq y^{**} \) choose to work only in the formal sector or to remain unemployed until they find a formal sector job. We determine \( y_r \) to match the agricultural worker share. Two cut-off parameters, \( y^* \) and \( y^{**} \), are endogenously determined in a way to best match our calibration targets.

We have four parameters \( \{\alpha, \delta, \lambda, m(\theta)\} \) that represent frictional urban labor markets. We fix the parameter for the informal sector job destruction rate \( (\delta) \) at the value in GPP (2008) to reduce the computational burden. In this case, we assume that the informal job destruction shocks come every 5 years on average. Since there are no empirical findings on the shape of the matching function for developing countries, we assume a standard CRS function from the labor search literature.
for developed countries. The function is given by

\[ m(\theta) = \zeta \theta^\rho. \]

We take \( m(\theta) \) as a parameter, set \( \zeta \) to be unity and use a standard value \( 1/2 \) for \( \rho \). The benchmark steady-state value of \( \theta \) is given by \( \theta = m(\theta)^2 \).

As discussed in GPP (2008), the migration cost \( M \) is given by a no-migration condition, and the cost of a vacancy \( C \) by the formal sector firm’s free-entry condition.

A subset of model parameters \( \{\alpha, \lambda, A_1, m(\theta)\} \) is determined within the model in a way to best match the employment shares, average formal sector wage\(^{11}\), and unemployment duration. In this calibration exercise, we have a system of non-linear equations to solve for the unknown parameters. Generally, it is difficult or impossible to analytically prove both existence and uniqueness of the solution. Our computational approach is to construct the loss function\(^{12}\) and choose the parameters that minimize the loss value.

In cases in which uniqueness is guaranteed, the minimization problem above finds the unique solution. If multiple solutions exist, then we must have a criterion to choose one of them. Since we use a simplex-based algorithm for the minimization problem, multiple solutions implies that the solution is not independent of the set of initial guessing values. Thus, our rule to choose the solution is to create grid lines of parameters for initial guessing values for the candidate solution. Some initial guessing values end up with non-convergence, and some others may converge. Our

\(^{11}\)Average labor incomes in both the agriculture and the informal sector are estimated from the IFLS.

\(^{12}\)In our case our loss function is the average of the square of the log difference between the data targets and corresponding model outcomes. The reason why we choose the log difference is that some targets have different units of measure.
choice rule is to have the best solution, or solution of the minimum loss value within the parameter space we imposed.\footnote{We bound $\alpha \in (0,10)$, $\lambda \in (0,5)$, $A_1 \in (1,5)$ and $m(\theta) \in (0,10)$ in our calibration. Thus, our restrictions are such that both the formal and informal sector job opportunities arrive on average 10 times at best, and the idiosyncratic shock hits formal sector jobs on average 5 times. The maximum formal sector TFP is constrained to be 5.}

### 2.4.1 Implications of Benchmark Parameterization

Benchmark parameters are listed in Table 10, and outcome variables are in Table 11. Notice that our calibration results are quite accurate in the sense that they are matching the targets almost perfectly.

Differences in sectoral productivities are translated into those in sectoral average incomes. The results suggest that the formal sector TFP is substantially higher than that in agriculture (e.g. $A_1/A_a = 8.7302$) while the ratio of the average incomes is $w/Y_a = 2.8833$. Additionally, informal sector job opportunities come more frequently than formal sector ones. Unemployed workers find about 1.5 formal job opportunities during each year, on average (e.g. $m(\theta) = 1.5281$), although they do not necessarily take them. They also meet 4 informal jobs on average (e.g. $\alpha = 4.0163$).

As with the rate of informal job arrival, the informal job destruction comes infrequently (e.g. $\delta = 0.1757$). On the other hand, idiosyncratic productivity shocks on formal jobs arrive once in every 3 years, on average (e.g. $\lambda = 0.3021$). In the steady state, a flow value in the formal sector is greater than that in the informal sector by 2.4%.

The cost of vacancy, $C$, is about 30% of an average formal sector wage. The cost of migration takes about 90% of average income in agriculture. For the other parameters, we use either standard
values in the literature or normalized ones.

In Figure 7, we draw the reservation productivity for formal sector workers. Since \( y_r < y^* \) as shown in Table 11, the level of skill for formal sector workers must be greater than \( y^* = 0.3719 \). The reservation productivity is flat between \( y^* \) and \( y^{**} \) and increases with \( y > y^{**} \). For workers with \( y \in [y^*, y^{**}] \), they have the similar outside option value of unemployment, so that the slope of the reservation productivity is flat. For workers with \( y > y^{**} \), the outside option value increases with \( y \). So does the reservation productivity. There is a kink point at \( y^{**} = 0.6015 \) on right hand side of which the slope of the graph becomes steeper.

Overall, our benchmark calibration shows, as expected, that the formal sector is the most productive and has a relatively high rate of job turbulence. Thus, the formal sector has relatively higher job flow rates\(^{14}\) than informal sector.

### 2.5 Outcome Validation

In this section, we validate our model by comparing our calibration outcomes with data observations of Indonesian labor markets in 1997. Thus, this validation exercise is made to the extent that the IFLS permits. We use the outcomes from these exercises partially to support the use of the model in characterizing labor markets in Indonesia and developing countries of the similar characteristics.

There are several outcomes we can use for this purpose. Specifically, we examine 1) skill distribution, 2) formal wage distribution, 3) structure of segmentation by skill, and 4) employ-

\(^{14}\)The job flow rate in the formal sector is 0.7, whereas the rate in the informal sector is 0.46.
2.5.1 Skill Distribution

Our parameterization assumes that the skill distribution is continuous and follows a $\beta$-distribution with two parameters estimated by a proxy variable years of education. The skill distribution is not the same as the distribution of education. The education variable is a proxy in a sense that labor market segmentation by skills follows closely on the segmentation by the level of education. Our result suggests that model distribution is smooth and unimodal, whereas the data distribution is jagged and multi-modal with the peak at 6 years (complete elementary school), 9 years (complete middle school), and 12 years (complete high school).

Figure 8 draws the skill distributions from the model and data respectively. The differences in the shapes come mainly from the continuous skill distribution in the model in contrast with discrete education data. However, the first two moments of the distributions are identical by construction. Although the simulated skill distribution does not match well with the empirical skill distribution, the simulated distribution remain useful for describing average behavior of workers in a multi-sector labor market.

---

15 We first normalize the proxy variable by the maximum years of schooling and then estimate two parameters of $\beta$-distribution in the support of $[0, 1]$.
16 The level of education may be one component included in measuring one’s skill
2.5.2 Wage Distribution

Given the level of skill, a formal sector wage is determined by Nash bargaining, and it can be expressed by a weighted average of a flow value of a worker-firm match and an outside option value of the worker’s unemployment.\textsuperscript{17} Since it is impossible or difficult to derive the analytic expression for the wage distribution, we numerically construct this distribution. Specifically, we generate employed workers in the formal sector from their skill distribution, and then we simulate a history of idiosyncratic skill shocks for each worker. The distribution of the idiosyncratic shocks is constructed from the Poisson distribution with the annual shock arrival rate of $\lambda$. As discussed in ANV (2009), we assume that the idiosyncratic shocks are independent across individuals and previous history of shocks.

Figure 9 illustrates the observed and simulated wage distributions. Clearly, the distribution in the data is more dispersed than the one simulated by model parameters. Additionally, the observed wage distribution has a long right tail, which is not a feature of the simulated distribution.

As the skill distribution is constructed by individual schooling observations, schooling distribution does not have an appropriate shape to match the higher moments of the wage distribution. In other words, we need to have some other factors to produce the wage dispersion not captured by the model. One way to improve this distributional discrepancy may be to use a better proxy for the skill. For example, we may construct the skill distribution by the partial predicted values from running Mincerian wage equation. In other words, we choose some of independent variables that is directly related to one’s skill and use the predicted values from only these variables as the proxy

\textsuperscript{17}See the section 1.3.4 in Chapter 1.
for the skill. However, since our objective is not to match the wage distribution itself but only the average wage, this discrepancy is not likely a serious flaw for our purposes.

### 2.5.3 Structure of Segmentation by Skill

There are two types of segmentation in labor markets. The first one is made by a geographical distance between urban and rural sectors. When workers in one location migrate into another, they must bear a migration cost, which works as one type of entry barriers. In equilibrium, workers in all locations do not have any incentive to migrate.

Another type of segmentation is made within the urban labor markets. It is not profitable for formal sector firms to hire sufficiently low skilled workers on one hand. The low skilled workers have no incentives to have a job in the formal sector when their values of the informal sector jobs are greater than those of formal sector ones, on the other hand. Therefore, there exist some low-skilled workers who choose to work in the informal sector even when formal sector opportunities arrive. Similarly sufficiently highly skilled workers choose only formal sector jobs.\(^\text{18}\)

The segmentation in our model is endogenously made by individual skills, which is the main modeling idea of ANV (2009) and GPP (2008). Model parameters that represent the segmentation are \(\{y_r, y^*, y^{**}\}\), where \(y_r\) is the threshold value of the skill in urban-rural segmentation; \(\{y^*, y^{**}\}\) determines formal-informal segmentation in the urban labor market.\(^\text{19}\) We list sectoral educational statistics in Table 12.

Average education levels across sectors can be ranked from agriculture to informal and to

---

\(^{18}\)See section 3.6 in GPP(2008).

\(^{19}\)See ANV (2009) and GPP (2008) for explanation in detail.
formal sector in ascending order, which qualitatively matches our model outcomes in Table 12. Our model can also reproduce the segmentation by skill observed in the data. Overall, our model is off by 2 years from the average years of education across sectors. Further, notice the average skill/education of the unemployed workers in both data and model. The average value is slightly lower than that of formal sector workers. Thus, our model outcome is in line with the view that unemployment is a 'luxury', since those high-skilled unemployed workers tend to work in formal sector and earn relatively high incomes.

2.5.4 Employed Skill Distribution

Figure 10 illustrates the share of workers in each sector across skill levels. The first plot shows the agricultural sector. Each circle represents the share of workers employed at a given level of skill. Roughly, 60% of workers with less than 3 years of education are employed in agriculture, and the percentage decreases as the level of skill increases. Our benchmark result implies that all workers with the education less than 5.6 years are employed only in agriculture. Qualitatively, the employment share distribution in agriculture is well squared with data observations in that relatively low-skilled workers are more likely to be employed in agriculture.

The employment share in the informal sector has an inverse-U shape in our model with respect to years of education, and it monotonically decreases after 5.6 years. Similarly, the share has an inverse-U shape in the data as shown in Figure 10. The formal employment share in our model is monotonically increasing with the level of skill, and so is the share in the data. Overall, Figure 10 shows that relatively, low-skilled workers are found in the rural agriculture in both the model
and the data; in urban sectors, the employment share in the informal (formal) sector decreases (increases) with the level of skill in both the model and the data.

We illustrate the cumulative distributions of skill across sectors from both the model and the data in Figure 11. First, the model suggests that most workers in the agricultural sector should have a level of education lower than 6 years. In the data, 80% of workers in the agriculture have education lower than 5.6 years. Our benchmark result implies that employed formal sector workers have the level of education above 5.9 years. This also well squares with the data. The data suggests that 88% of workers in the formal sector have education above 5.9 years.

The second plot in Figure 11 illustrates the distribution of skills in the informal sector. Our model shows a more concentrated profile of skills in the informal sector than the data draws. In particular, the model predicts few workers with education below 5.6 years. However, the data shows that 31% of workers in the sector have a level of education below this level.

Overall, the level of education is a plausible proxy for the skill distribution in a sense that we match the first moment of the sectoral distributions in our calibration. However, the education proxy may not be sufficient if our calibration targets higher moments of the distributions.

2.5.5 Unemployment Rates by Skill

Workers in the rural sector are always employed in subsistence agriculture, which acts as employment of the last resort. As such, the unemployment rate of those low-skilled agricultural workers should be low. Unemployment is an urban phenomenon, and it is considered as a 'luxury' that only highly skilled workers can afford to, and unemployment is only worthwhile if a good employment
opportunity is pending. This implies that unemployment rate should rise with education.

Figure 12 illustrates unemployment rates by the level of skill. The observed unemployment share by level of education is increasing up to 12 years (which amount to 92% of the workforce). Unemployment shares for those with education less than 7 years are below 7%. For moderate skilled workers, unemployment shares are between 7% and 14%. However, for workers with more than 12 years of education, the unemployment share decreases again, suggesting that the opportunity cost of unemployment for high-skilled workers is very high. The model is not able to replicate this low-unemployment share of the high-skilled workers. However, this should not be a huge concern since these workers represent less than 7% of total workforce. Notice that the unemployment share is constant for workers with \( y > y^{**} \). High-skilled workers are more vulnerable to unemployment as their reservation productivity increases in \( y \) for \( y > y^{**} \). However, the difference between one’s potential productivity/skill \( y \) and reservation productivity \( R(y) \) becomes wider as \( y \) increases for \( y > y^{**} \), so that workers become less vulnerable to unemployment. In equilibrium, two effects offset each other, and unemployment share become steady for \( y > y^{**} \).

### 2.5.6 Summary

The model assumes segmented labor markets: a rural agriculture sector, and informal and formal urban sectors. These sectors have distinct wage-setting mechanisms and production technologies. Given these sectors, workers, indexed by skills, endogenously choose their sector of work. When we use the educational attainments as a proxy for skill, our model closely matches an average skill distribution across sectors. Also, in our model, high-skilled workers are found in the productive
formal sector, whereas low-skilled workers are engaged in production at either agricultural or informal sector. This segmentation by skill is also verified in the data. Further, the unemployment shares increase with the skill, which is also observed in the data. However, our model fails to match low unemployment shares for high-skilled workers.

As for the income distribution across sectors, all workers in the agriculture and informal sectors are assumed to earn average income in the data so that our model matches the data means by construction. Our model generates a wage distribution in the formal sector. Although the model matches the average wage, its distribution is clustered around the mean, whereas the data distribution is left-skewed. Overall, we think that the model is well equipped to explain average behavior of workers in the multi-sectoral labor markets.

2.6 Conclusion

We take a view of segmented labor markets in developing countries and choose a multi-sector labor search and matching model of GPP (2008). First, we validate our use of the model by comparing model outcomes with data observations. We show that skill-sector matching patterns are compatible between the model and data observation. In other words, high-skilled/educated workers tend to work in the formal sector, and relatively low-skilled/educated workers in the informal or agriculture, in both the model and the data.

In summary, the model turns out to be well equipped to explain labor market structure. Given this validation of the model, we may push a step further into implementing some policy instruments.
Chapter 3

Quantitative Study of Cross-Country Intergenerational Mobility

3.1 Introduction

Empirical studies have measured intergenerational labor income\(^1\) mobility\(^2\) in the U.S. and in some other countries. These studies’ results are important because they have strong implications for understanding individual economic success. One’s standing in the income distribution might be a result of individual effort (i.e., work habits) or acquired ability (i.e., human capital or I.Q.). However, if those factors are heavily influenced by family background (i.e., family income, social class and family reputation), one’s standing is at least partially deterministic and moreover, may

---

\(^1\)Hereafter, we simply use ‘income’ to denote labor earnings.

\(^2\)Intergenerational elasticity as a measure of mobility ranges from 0.3 to 0.6 in the U.S. depending on the data source and definition. Solon (1999), D’addio (2007), Björklund and Jäntti (2009) provide surveys of this literature.
be sustained throughout generations.

Thus, it is important to study the sources of social standing and the channels through which social standing is transmitted intergenerationally.\textsuperscript{3} In this paper, we modify Solon’s (2004) specification of Becker and Tomes’s (BT) (1992) model and study the quantitative importance of the sources of cross-country differences in intergenerational income mobility.

As in BT, we consider two factors; namely, human capital investment and ability endowment. We define human capital as a factor of the income-generating function. Like the BT model, altruistic parents provide their children with many learning opportunities so that children may be equipped with as high an earnings capacity as possible later in their life. The parental resource allocation for those opportunities is called the human capital investment, and a child’s earnings capacity reflects his/her human capital holdings.

Thus, the size of the per-unit returns on human capital investments is quantitatively important for intergenerational mobility. If the returns are relatively large, income-rich altruistic parents invest more in their children; therefore, income mobility is relatively low. In other words, one’s standing in terms of the income distribution tends to persist throughout generations.

We consider learning ability as the second source of intergenerational income mobility. Ability is often decomposed into cognitive and non-cognitive components. Here, we do not make any distinction between the two types but rather simply use the term ”learning ability”. Learning ability is understood as an investment-augmenting technology. When two persons have the same level

\textsuperscript{3}D’addio (2007), Björklund and Jäntti (2009), Altonji and Dunn (1996), and Corcoran et al.(1992) provide empirical analyses of this question. For theoretical references, see Ben-Porath (1967), Becker and Tomes (1992), Han and Mulligan (1997), Solon (2004) and Huggett et al.(2006).
of human capital investment, but their learning abilities are different, the one who has a greater learning ability can produce a higher level of human capital. Thus, learning ability augments human capital investment. In other words, the efficiency level of the human capital investment is higher for someone with a greater learning ability. We assume that learning ability is exogenously transmitted and follows an AR(1) process.\(^4\)

Given these channels of mobility, we calibrate the model for 10 OECD countries and discuss cross-country implications. Specifically, we study how institutional differences represented by different parameterizations are related to intergenerational mobility. We then analyze the importance of ability transmission and human capital investment with respect to mobility. Finally, we focus on the impact of a direct governmental subsidy on human capital investment.

Our results are as follows. Mobility in countries with higher returns on the human capital investment tends to be more intergenerationally related so that one’s standing in the income distribution depends more on parental income and therefore on the parental human capital investment. Second, countries with high variation in learning ability transmission tend to generate high cross-sectional inequality in terms of income distribution. When we focus on how the degree of intergenerational ability variation affects mobility, it appears that the magnitude of heterogeneity is not critical in determining mobility. However, when each child is given relatively equal opportunities in human capital accumulation or provided with relatively homogeneous human capital

\(^4\)This specification is generally adopted within the theoretical literature. The empirical literature has shown that learning ability as measured by cognitive ability is intergenerationally correlated and has a significant impact on market outcomes. This correlation ranges from 0.42 to 0.72 (Daniels et al., 1997, Bowles et al. 2003). The direct effect of learning ability on market outcomes is 0.15 (Bowles et al. 2003)
investments, intergenerational mobility is shown to improve significantly.

Our main contribution to the literature is to quantitatively study the BT model and compare intergenerational human capital transmission functions across countries. The remainder of the paper proceeds as follows. In Section 2, we describe the dynasty model of human capital investment. In Section 3, we discuss our benchmark parameterization. In Section 4, we apply the proposed model to 10 OECD countries and discuss outcomes that explain cross-country intergenerational mobility patterns. In Sections 5 and 6, we analyze how learning ability variation and direct governmental human capital investment subsidies affect intergenerational mobility. We conclude the paper in Section 7.

3.2 Literature

Solon (1992) and Zimmerman (1992) challenged BT’s empirical claim that intergenerational income mobility in the U.S. is no worse than in European countries. This challenge involved methodological issues related to life-time income. When life-time income is proxied by the average or present discounted income for a short time span, Solon (1992) showed that there is an estimation bias for mobility. Solon (1992) adopted the error-in-variable method to construct life-time income and found that the U.S. is no better than European countries in terms of intergenerational mobility.

These authors’ method of measuring intergenerational income elasticity (or correlation) has been widely adopted in country studies.\(^5\) Different sample definitions and time periods make direct international comparisons difficult. Furthermore, estimation methods differ depending on

\(^5\)Björklund et al. (2009) and Solon (2002) provide surveys of such international comparisons.
the availability of data. Thus, studies on international income mobility tend to restrict themselves to the comparison of the regression coefficients of simple linear regression models.

Björklund et al. (2009) and Solon (2002) showed that the coefficient of the standard linear model is around 0.2 for Nordic countries (Denmark, Sweden, Finland and Norway), whereas the coefficient for France, Italy and the U.S. is around 0.4. Using a similar sample, time period and methodology, Björklund et al. (2000) estimated intergenerational income elasticity for the U.S. and Sweden and showed that the elasticities differ significantly.

These cross-country studies naturally lead us to question what underlying factors cause such differences. One potential explanation frequently mentioned in the literature is the heterogeneity of institutions among countries. In our paper, we take BT’s model of human capital investment and parameterize it for the heterogeneity of institutions.

Our model also builds on the work of Hassler et al. (2002) and Restuccia et al. (2004). Both studies showed that educational investment (i.e., public education versus productivity growth in Hassler et al. (2002) and early childhood education versus college education in Restuccia et al. (2004)) affect both intergenerational income persistence and cross-sectional income inequality. These studies utilized dynamic general equilibrium models with acquired ability and parental income heterogeneity. The implication of Hassler et al. (2002) was that productivity differences (i.e., TFP), rather than public education, cause European countries to be more equal and less mobile than the U.S. Restuccia et al. (2004) emphasized the role of early educational investment in intergenerational mobility and concluded that college education is mainly responsible for cross-sectional inequality.
3.3 Model

We use a two-period overlapping generations model. There is one child and one parent in each generation. The parents in each generation are assumed to act altruistically toward their immediate offspring. Because a child is assumed to act altruistically toward his/her future child, a given parent is essentially altruistic toward all his/her future descendants. This assumption justifies the dynastic structure of preference because a parent’s actions toward the immediate child ultimately affect the utility of all descendants thereafter.

The parent at time $j$ earns income $y(h_j)$ from his/her human capital $h_j$. Given disposable income $y^d(h_j)$, the parent decides how much the family will consume $c_j$ and how much to invest $s_{j+1}$ in the child’s human capital accumulation. The learning ability $a_{j+1}$ of the child at time $j+1$ is transmitted partly from the parent’s $a_j$ and partly by a random environmental effect $\varepsilon_{j+1}$. Specifically, we assume that learning ability is an AR(1) process.6

Given the structure of the model described above, our problem is characterized by

$$U(y_1; a_1) = \max_{\{c_j, s_{j+1}\}_{j=1}^{\infty}} E[\sum_{j=1}^{\infty} \beta^{j-1} u(c_j)]$$

subject to

$$c_j + s_{j+1} = y^d_j$$

$$\log(a_{j+1}) = \rho \log(a_j) + \varepsilon_{j+1}, \quad \rho \in (0, 1)$$

$$h_{j+1} = a_{j+1}s^\gamma_{j+1}, \quad \gamma \in (0, 1)$$

---

6We use the specification from Becker and Tomes (1986) and Solon (2004).
\[ y^d_{j+1} = \bar{\kappa} h^p_{j+1} \]
\[ \varepsilon \sim i.i.d.N(0, \sigma^2_v) \]

Given \( y_1, a_1 \)

We assume that there is no borrowing or lending, as in Solon (2004). The parent makes a human capital investment \( s_{j+1} \) in the child. When the parent retires, no debt is transmitted to subsequent descendants. Thus, life-time disposable income determines the human capital investment.

The human capital production function is assumed to be concave in the investment \( s_{j+1} \). This concavity captures the diminishing marginal absorption rate of the investment at any given \( a_{j+1} \). When opportunities for human capital accumulation are relatively scarce, the child can better take advantage of those opportunities by generating relatively larger marginal returns. However, when human capital accumulation is relatively abundant, the absorptive intensity is smaller because there are many other alternatives.\(^7\) However, the return is critically affected by the magnitude of \( a_{j+1} \), so the return might be larger for a child from an income-rich family with high learning ability than a child from a poor family with low learning ability.

Income is generated from the combination of common non-human capital factors (\( \bar{\kappa} \)) and the human capital accumulated (\( h_{j+1} \)) accumulated by an individual. For technical simplicity, we assume that the rental price of human capital is \( p = 1 \), so (disposable) income exhibits diminishing returns on human capital investment (i.e., \( y^d_{j+1} = \bar{\kappa} a_{j+1} s^2_{j+1} \)).

\(^7\)This concavity assumption supports one popular idea in the policy debate over human capital investment, namely, that the government should focus on children from poor families because of their relatively larger marginal returns on investment. This claim, however, may not be true according to the above model if children from high income families also tend to have higher learning abilities.
In each period, $j = 1, 2, \ldots$, the family faces the same optimization problem with different state values. We can therefore reformulate the optimization problem as a dynamic programming problem as follows.

$$
V(y, a; \Theta) = \max \{ u(c) + \beta E[V(y', a'; \Theta)|a] \} 
$$

subject to

$$
c + s' = y^d 
$$

$$
\log(a') = \rho \log(a) + \epsilon, \quad \rho \in (0, 1) 
$$

$$
h' = a' h' \quad \gamma \in (0, 1) 
$$

$$
y' = \bar{\kappa} h' 
$$

$$
\epsilon \sim N(0, \sigma_\epsilon^2) 
$$

$$
\Theta = \{ \beta, \tau, \rho, \gamma, \sigma_\epsilon^2 \} 
$$

This model has a similar structure to the standard neo-classical growth model. We define state variables as $(y, a)$ instead of $(s, a)$. This specification is made to utilize the empirical distribution of lifetime income statistics. Both value and policy functions are proven to have all the usual features.\(^8\) In equilibrium, there exists a bivariate stationary distribution $(y, a)$.

\(^8\) $V(y, a)$ is continuous, increasing in $h$ and $a$, and is concave in $y$, and $y(y, a)$ is single-valued. Refer to Stokey et al. (1989), especially, chapters 4 and 10.
3.3.1 Discussion

In some BT-type models, resources are transmitted intergenerationally partly through indirect human capital investment and partly through direct pecuniary bequest\(^9\) assuming a perfect capital market. These models may be theoretically and/or quantitatively useful in characterizing the importance of the channels through which an efficient investment is made. However, our purpose is to investigate the role of human capital production efficiency in explaining intergenerational correlation. The bequest is therefore not essential for our purposes.\(^{10}\)

As in Solon (2004), we do not specify a borrowing constraint for human capital investment.\(^{11}\) Most of the cross-country intergenerational elasticities or correlations are estimated without controlling for the effect of the constraint. Therefore, the parameters of the model calibrated with a credit constraint would not be consistent with empirical estimates. Solon’s (2004) parsimonious re-characterization of the BT model without a credit constraint and physical capital serves our purposes of quantifying the efficiency of human capital production while being consistent with empirical estimates of intergenerational statistics.\(^{12}\)

\(^9\)A child’s learning ability plays a key role in this decision margin such that the parent tends to spend more on human capital investment if the child is of high learning ability; otherwise, more is spent on the bequest.

\(^{10}\)In other words, we are more interested in the process of human capital generation than decision problems related to parents’ efficient investments.

\(^{11}\)In their simulation study of the BT model, Mulligan and Han (1997) showed that the effect of a borrowing constraint is diluted by heterogeneity in individual characteristics (i.e., parental income and learning ability). Grawe (2004) examined the credit constraint hypothesis and found that the constraint is not the driver of a non-linear intergenerational income curve.

\(^{12}\)Solon (2004) reformulated the BT model to study the relationship between cross-sectional income inequality and intergenerational income persistence without the specific objective of making an international comparison.
We can formulate the human capital generating function as follows.

\[ h_{j+1} = a_{j+1} (s_{j+1} + \bar{s})^\gamma \quad \gamma \in (0, 1), \]

where \( \bar{s} \) is the size of the public human capital subsidy. This specification provides some dimensions with which to experiment on the influence of a public human capital subsidy on intergenerational mobility. However, it is not clear why both types of investments are additively specified, especially because the subsidy has the characteristics of a public good. In our model, this public investment is incorporated multiplicatively and captured by the parameter \( \bar{\kappa} \).

### 3.4 Benchmark Parameterizations

The period utility function is specified by

\[ u(c) = \log(c) \]

This homothetic utility function is given so that households make a positive human capital investment within their budget in equilibrium.

We include a set of parameters \( \{\beta, \gamma, \bar{\kappa}, \rho, \sigma_v\} \) in our model. Because \( \bar{\kappa} \) does not affect the intergenerational log-income correlation, we can choose any arbitrary value, and we fix it at 5. We use \( \beta = \left( \frac{1}{1 - 0.04} \right)^{20} \) so that one generation indicates 20 years. We then choose the other three parameters to best match the three data moments for the U.S. Target statistics are listed in Table 14.

The intergenerational elasticity of log earnings in the U.S. ranges from 0.3 to 0.6. We choose 0.5 as the target. We adopt the log-income variance of 0.36 from Restuccia et al. (2004). Restuccia et al. (2004) computed the value from the data set used by Mulligan (1997). The Gini is adopted
from Björklund and Jäntti (2009) and was derived from the Luxembourg Income Study (LIS) wave 1. Because we have a stationary log income distribution in our model, intergenerational income elasticity is the same as intergenerational correlation.

We run our calibration assuming that one particular set of target statistics prevails in the long run stationary state. Benchmark parameter values are listed in Table 13.

The standard deviation $\sigma_v$ in the learning ability process is responsible for most of the log-income variance because it affects the distribution of learning ability among individuals.\(^{13}\) The parameter $\gamma$ of investment elasticity controls the degree of intergenerational income correlation. The child’s log-income function implies that more of the parental human capital investment can be transmitted with an increase in $\gamma$ because human capitalization increases (i.e., $h = as^\gamma \leq as^\gamma' = h'$ if $\gamma \leq \gamma'$). Therefore, an income-rich altruistic parent has a strong incentive to make a large human capital investment when human capital production elasticity ($\gamma$) is high, which in turn increases intergenerational income correlation.

The correlation coefficient parameter $\rho$ for the learning ability process may significantly affect the intergenerational income correlation. Because income-rich parents tend to be endowed with a high learning ability, their children may take advantage of the transmission of intergenerational learning ability and of the high human capital investment from their rich parents. Parameters $\sigma_v$, $\rho$ and $\gamma$ are important for the Gini coefficient because they affect the cross-sectional income

\(^{13}\)It is clear that cross-sectional income variation is also affected by heterogeneous human capital investment. The contribution is, however, reduced through the human capital production elasticity parameter $\gamma \in (0, 1)$. It turns out that heterogeneity in learning ability is quantitatively more important than that in human capital investment. We will discuss this point in greater detail in the next section.
distribution. However, the magnitude is ambiguous.\textsuperscript{14}

The three parameters \(\{\gamma, \rho, \sigma_v\}\) are simultaneously determined within our model in a way that best matches the statistics in Table 14. In choosing \(\{\rho, \sigma_v\}\), we choose the stationary learning ability distribution. As shown in Table 14, our benchmark specification matches target statistics at the tolerance\textsuperscript{15} level \(10^{-4}\).

Suppose the human capital investment is made by a planner such that all children are provided with the same amount of human capital investment. Cross-sectional income variation in this case is determined solely by stationary ability variance\textsuperscript{16}, and the intergenerational income correlation is equivalent to the intergenerational ability correlation.\textsuperscript{17} Because the intergenerational transmission of human capital is only made through the transmission of exogenous ability, we may interpret \(\rho\) as a lower bound for the intergenerational correlation. Further, in our benchmark specification, the direct effect of learning ability correlation explains 20\% of the intergenerational correlation for the U.S. Due to the construction of the model, the intergenerational log-income correlation is explained more by the ability correlation when countries are more intergenerationally mobile.

### 3.5 Cross-country Comparison

We calibrate our model for 10 OECD countries. Countries are represented using different parameterizations of the model. Specifically, they are described by parameters that characterize human

\textsuperscript{14}In fact, all three parameters simultaneously affect the targets shown in Table 14 to some degree.

\textsuperscript{15}We use our criterion function as the average square of the log difference between the model outcomes and the data targets.

\textsuperscript{16}Note that \(\log(y) = \log(\bar{K}) + \log(a), \text{var}(\log(y)) = \text{var}(\log(a))\).

\textsuperscript{17}\(E(\log(y') \log(y)) = E(\log(a') \log(a)) = \rho E(\log(a)^2)\). Hence, \(corr(\log(y'), \log(y)) = \rho\).
capital production efficiency and the learning ability distribution. In this section, we study how institutional differences affect the cross-country intergenerational log income correlation as well as the cross-sectional income Gini.

We assume that the correlation of intergenerational learning abilities is invariant across countries. First, individuals from all countries are, on average, subject to the same degree of genetic transmission. Second, given that all parents in each country are assumed to act altruistically toward their children, the average exogenous environmental effect that potentially affects learning ability should be similar across individuals in all countries; for example, the effect of social class, reputation and positive attitudes of parents on children’s learning ability formation may be captured by the correlation coefficient. Thus, the assumption of the invariant correlation of intergenerational learning abilities may not be too restrictive; therefore, we take $\rho$ from the U.S. as the benchmark value for all countries we consider.

Most countries do not have sufficiently long panel data that cover the complete life-cycle income series for at least two generations.\textsuperscript{18} Although we do not have descriptive statistics for lifetime income, we do have empirical estimates of intergenerational income elasticity and a cross-sectional inequality measure (that is, the Gini coefficient).\textsuperscript{19} We assume these two data statistics prevail in the long-run steady state and calibrate the remaining two parameters $\{\sigma, \gamma\}$.

\textsuperscript{18}The intergenerational income correlation in the empirical literature is generally computed using the proxy lifetime income variable (i.e., the time average or present discount value of several years of income).

\textsuperscript{19}For our data moments, we use the Table Y in Björklund and Jäntti (2009).
3.5.1 Summary of Results

We tabulate the data and model moments, and parameter values in Table 15. Notice that countries with a high standard deviation $\sigma_v$ in the ability process unambiguously have high log-income Gini values as shown in Figure 13, whereas the standard deviation has a weaker relationship with intergenerational income persistence. For example, Denmark has a relatively low value of intergenerational correlation at 0.123 and a relatively high value of $\sigma_v$ equal to 0.473. On the other extreme, the U.K. has a relatively high correlation value at 0.5, with $\sigma_v$ equal to 0.4.

Intuitively, as learning ability variance increases, so does the variance of the human capital holdings. Thus, cross-sectional income inequality, measured by the Gini index, tends to increase. Intergenerational persistence, however, is significantly affected by country-specific human capital production efficiency, which is measured by parameter $\gamma$ in our model. The high value of the efficiency parameter together with random effects can amplify intergenerational income persistence, as in the U.K. However, a low efficiency parameter with the same degree of learning ability can suppress intergenerational income persistence, as in Denmark.\(^{20}\)

Countries with a high intergenerational correlation have large marginal returns on the per-unit human capital investment. As shown in Figure 14, the intergenerational income correlation monotonically increases with the efficiency parameter $\gamma$. Consider a country with a high level of human capital production efficiency (i.e., a high $\gamma$). Children from poor families tend to be less endowed with learning abilities because their parents also have a low level of ability on average. Given their low incomes, a relatively small human capital investment in these families would be

\(^{20}\)See Table 15
optimal. Alternatively, children from rich families tend to be endowed from their parents, who presumably have a high learning ability, with the same high ability. Thus, rich altruistic parents have incentives to make relatively large human capital investments. Hence, in high-γ societies, intergenerational income persistence tends to be large, as Figure 14 shows.

As for the impact of the efficiency parameter γ on the cross-sectional income Gini coefficient, we do not observe much variation in Gini across γ. In other words, the differences in human capital production efficiency are not quantitatively important in determining cross-sectional inequality.

Why do countries have different levels of human capital production efficiency? In this model, we assume that there is no direct governmental human capital investment subsidy to households. However, we assume that the government supports such investments through public education expenditures and various tax rebates on education spending. These types of governmental policies could be viewed as public human capital investments in the context of our model. Thus, a country with large public investments could have low marginal returns on private human capital investment on average, which in turn translates into a low value of the efficiency parameter γ in our model. Using OECD education spending time-series data, we display the relationship between intergenerational correlation estimates and the ratio of government education spending to GDP in Figure 15.

Except for Canada, countries with a high ratio of public education spending to GDP tend to have a low intergenerational correlation (a low γ21). This figure may provide some implications for both the quantitative and qualitative enhancement of public education as a tool for a more intergenerationally mobile society.

21See Figure 14.
3.6 Ability Distribution on Intergenerational Mobility

In the previous section, we showed that the learning ability variance $\sigma_v^2$ has a quantitatively weaker relationship with the intergenerational income correlation than the efficiency parameter $\gamma$. However, as shown in Table 15, the overall variance is positively related to the intergenerational correlation. Thus, in this section, we focus on how changes in variance affect the degree of mobility.

When $\sigma_v$ is low, the variance of learning ability is low. Therefore, individuals are relatively homogeneous in their learning ability. In this case, parents’ human capital investment is the main determinant of children’s earnings capacities. We may therefore expect a high covariance among intergenerational income but a low variance among cross-sectional income, which yields a low value for the intergenerational income correlation.

However, when the variation of learning ability is relatively large, the relative income of offspring in the cross-sectional distribution may significantly differ from that of parents, whereas the variance of cross-sectional income is large. In this case, we expect a low intergenerational correlation with a low intergenerational covariance and a high cross-sectional variance.

From the benchmark parameterization, we show the intergenerational correlation pattern with respect to $\sigma_v$ in Figure 16. The correlation increases slightly when $\sigma_v$ increases from 0.1 to 0.3, and it decreases thereafter. As expected, when learning abilities are relatively homogeneous (i.e., $\sigma_v$ up to 0.4), the intergenerational covariance effect dominates, and the correlation is relatively large. As abilities become more heterogeneous (i.e., $\sigma_v$ greater than 0.4), the cross-sectional variance effect dominates, and the correlation decreases.

How important is the influence of the ability distribution on the intergenerational log-income
correlation? For $\sigma_v \in [0.1, 0.7]$, the intergenerational correlation is within $[0.45, 0.54]$. $\sigma_v = 0.7$ maps onto the stationary log-income variance of 0.8, whereas 0.1 of $\sigma_v$ maps onto a log-income variance at 0.0017. Although an increase in $\sigma_v$ from 0.1 to 0.7 drastically alters the stationary cross-sectional log-income variance, $corr(\log(y'), \log(y))$ changes only slightly, from 0.45 to 0.54.

Consider an economy in which the variance of the learning ability distribution increases. When it is relatively small (i.e., a value of up to 0.4 for $\sigma_v$), human capital investment is relatively less risky in its returns. Thus, income-rich (income-poor) parents of high (low) ability tend to transmit high (low) human capital together along with their high (low) ability. In this case, a relatively high intergenerational income correlation is expected. However, when the variance becomes sufficiently great (i.e., greater than 0.4 of $\sigma_v$), human capital investment becomes riskier because the returns become more contingent on the randomness of error in the learning ability process. Thus, in this case, the intergenerational income correlation decreases.

3.7 The Effect of Direct Governmental Human Capital Investment on Intergenerational Mobility

In Section 3.5, we made the plausible argument that cross-country heterogeneity in public education spending may be closely related to the magnitude of the efficiency parameter $\gamma$ in the human capital production function. In this section, we ask how intergenerational mobility would change if individuals became more homogeneous in human capital investment. We consider a direct governmental subsidy to human capital investment. Specifically, we focus on the case in which a

---

22 The stationary variance of cross-sectional log-income is 0.36 in our benchmark.
part of income tax revenue is used to directly subsidize human capital investment in a lump-sum manner.

\[ y' = \bar{\kappa}a'(s + \bar{g})^\gamma \]

Once the direct governmental subsidy \( \bar{g} \) is added in the human capital production function above, each family revises its human capital investment decision depending on its state \((y, a)\), which in turn influences the intergenerational income correlation. We show the correlation with respect to the level of subsidy in Figure 17. When the subsidy \( \bar{g} \) increases, the correlation decreases continuously.

A subsidy up to around 3% of the mean lifetime income does not change the overall investment behavior of households or the correlation. However, the correlation declines steeply with a sizable increase in the subsidy of more than 3%. Namely, all households decrease their human capital investment as the subsidy increases. The magnitude of the decline is relatively larger (smaller) if a household is income-rich (income-poor) because the marginal returns from the human capital investment are relatively smaller (larger) for income-rich (income-poor) families. With a sufficiently large subsidy, the role of human capital investment becomes negligible, even in income-poor families. Hence, the correlation declines at a slower pace and attains the minimum bound 0.1 of the correlation coefficient of the intergenerational learning ability.

Although direct subsidy can reduce the intergenerational log-income correlation, this policy may not be socially desirable. Clearly, our model is not fit to answer this type of question without

\[ \text{Notice that the purpose of this exercise is to see the effect of equalizing human capital investment on intergenerational mobility. The government’s budget is therefore not an issue here.} \]

\[ \text{Using the Euler equation, we can easily check that } \frac{d\ell}{d\bar{g}} < 0 \]

78
a plausible measure of the welfare impact of the change in human capital policy. Nevertheless, this direct subsidy exercise shows that the intergenerational income correlation marginally declines as human capital investment equalizes.

### 3.8 Conclusion

In this paper, we quantify a dynasty version of the BT model and study its implications. When we apply our model to 10 OECD countries, we find that cross-sectional inequality depends more on the exogenous ability transmission process than on the human capital investment. However, intergenerational mobility is heavily influenced by human capital production efficiency. Therefore, countries with low human capital returns on investment tend to be more intergenerationally mobile.

Given the simple structure of the model, we find that direct human capital investment subsidies are quite effective in reducing the intergenerational income correlation. For instance, a subsidy of 10% of mean income reduces the correlation from 0.51 to 0.2, whereas a subsidy of 20% reaches the minimum bound of 0.1. Hence, intergenerational income links can be substantially reduced by subsidy policies or similar measures.

The model in our paper does not have a general equilibrium structure, so it is not well suited to analyze the welfare consequences of subsidizing household human capital investment. However, the intergenerational income correlation marginally decreases with increasing subsidies as Figure 17 shows, so it would be interesting to extend our model to incorporate the government’s budget to assess the optimal subsidy level that would maximize social mobility, and to determine some criteria for social welfare.
Finally, one policy implication of the model is that the government should subsidize children from poor families so they may have more opportunities to be exposed to environments conducive to learning, which would enhance their chances of economic success. This type of normative statement may be justified in view of the BT strain of human capital investment theory because the marginal returns from subsidies are relatively larger for children from poor families. However, we do not yet know the extent to which the redistribution of resources through a subsidy is welfare-maximizing.
Bibliography


Appendix A: Definitions of Unemployment, Informality and Formality

How do we assign workers in the data to the four labor market states? We divide workers from the EMNV into several occupational categories: wage workers in the formal sector, wage workers in the informal sector, family enterprise workers, and the individually self-employed and employers, distinguishing between those workers in primary sector activities (agriculture, hunting, forestry and fishing) and those outside of primary sector activities. For each of these categories, we compute average earnings, poverty rates and the share of low earners among each category.

We attempt to find categories of workers whose earnings are distributed at the low quantiles and high poverty incidence rates. For Nicaragua, this method assigns family enterprise workers, individually self-employed and informal wage workers outside of agriculture to the low productivity informal sector. Formal wage workers are assigned to the high productivity sector, and agricultural workers (not agricultural employers) are all assigned to the agricultural sector. Employers in both agriculture and non-agriculture are left out all together. We believe that they are better understood as firms, rather than workers.

We can define unemployment in two different ways. The first is to use a narrow definition of

---

25 These states are unemployment, employment in the urban formal and informal sectors, and employment in the agricultural sector.

26 Employers are defined as self-employed workers who employ at least one paid non-family member at work. The individually self-employed are defined as those who are self-employed and do not work with other family members and do not employ other workers, and family enterprise workers are defined as the self-employed who work with other family members, whether paid or unpaid, in the same enterprise and do hire non-family workers.
unemployment in which only those actively seeking jobs are classified as unemployed workers. However, it is well known that the use of this definition is limited for developing countries. As labor markets are not well developed, job search mechanisms are mostly informal and are unlikely to be captured in the traditional labor force questionnaires. It is common in developing countries that a large number of non-employed workers who are not actively searching for a job would take a job if offered, but these non-employed workers are classified as "discouraged" in the standard definition. In addition, many inactive workers are temporarily inactive, waiting for harvests or other seasonal activities. These features often merit a "broader" definition of unemployment that takes into account discouraged and temporarily inactive workers.

Using the household survey, we classify workers as unemployed if they are temporarily inactive, discouraged or inactive for other reasons, and we look at the transition rates from these states to employment. We find that people classified as discouraged or as temporarily inactive have transition rates into and out of employment similar to those classified as unemployed under the narrow definition. This suggests that those temporarily inactive and those classified as discouraged workers are not very different from unemployed workers and as such, should be counted among the unemployed. This evidence suggests that for Nicaragua, the broad definition of unemployment better represents the status of the idle labor force. While the narrow unemployment rates are around 4%, the broad rate is 12%, as shown in Table 1.
Appendix B: Calibration

Step 1: Calibration of the Distribution of Skill and Idiosyncratic Shocks to the Formal Sector

In the first step, we specify the parameters of the shock process \(G(y)\) and those of the distribution of skills \(F(y)\). We assume that the shock process is described by a uniform distribution with a support of \([0, y]\). In other words, when a shock arrives, a new level of productivity of a worker of type \(y\) lies between 0 and \(y\), with an equal probability of ending up anywhere in this support. The uniform distribution is completely specified by its support.

We assume that workers’ skill is distributed according to a \(\beta\)-distribution \(F(y)\), and we proxy the skill by the level of education normalized by its maximum level. The \(\beta\)-distribution is flexible, with a finite support between \([0, 1]\). It can display a broad range of shapes depending on its two parameters \(a_{\text{beta}}\) and \(b_{\text{beta}}\). It can be skewed or have the standard bell shape.

Let \(a_{\text{beta}}\) and \(b_{\text{beta}}\) denote two parameters. The mean and variance of the \(\beta\)-distribution are given by

\[
E[y] = \frac{a_{\text{beta}}}{a_{\text{beta}} + b_{\text{beta}}} \\
var[y] = \frac{a_{\text{beta}}}{(a_{\text{beta}} + b_{\text{beta}})^2} \left(\frac{a_{\text{beta}}}{a_{\text{beta}} + b_{\text{beta}}} + 1\right)
\]

We use the mean years of education and its variance from the EMNV survey to calibrate the two parameters in the distribution of skill. Since our distribution has the support of \([0, 1]\), we normalize the distribution by dividing individual years of education by the highest level displayed in the data. The mean years of education among the active working age population in Nicaragua
is 4.59 years, with a standard deviation of 4.3 years. The maximum level of education observed among those in the labor force is 19 years. This means that we set

$$\frac{a_{\beta}}{a_{\beta} + b_{\beta}} = 0.302,$$

$$\frac{a_{\beta}}{(a_{\beta} + b_{\beta})^2(a_{\beta} + b_{\beta} + 1)} = 0.0529,$$

where 0.302 and 0.529 are the mean and variance of the normalized level of education.

**Step 2: Calibration of the Parameters Determining Earnings in the Informal Sector and in Agriculture.**

The earnings in agriculture are given by the following average product of labor:

$$y_a = A_{a}^\gamma L_a^{-1}.$$  \hspace{1cm} (1)

We estimate the labor elasticity ($\gamma$) with longitudinal data for agricultural production and employment from 1961 to 2005. Data on agricultural output ($Y_a$) come from the Central Bank of Nicaragua. To have a long enough series of agricultural employment, we use the household survey (EMNV) from 2001 and 2005 as well as the data from the FAO for the working age population in the agricultural sector. Using these data, we estimate that

$$\log Y_a = c + \gamma \log L_a.$$

The estimate of $\gamma$ is 0.628. Because income from agriculture is given by (1), we can compute $A_a$. Using the household survey, we calculate average income in the agricultural sector, ($y_a$), and
the share of workers in agriculture, \( l_a \). The average income in agriculture is $10,717 córdobas, and \( l_a \) is 27.1\%, as shown in Table 1. This allows us to compute \( A_a \).

Finally, we find the average income in the informal sector \( (y_0 = 12,519 \text{ córdobas}) \).

**Step 3: Calibration of Other Parameters in the Model**

We calibrate our model to determine the remaining parameters by matching the share of workers in the formal and informal sectors \( (N_0 \text{ and } N_1, \text{ respectively}) \), the share of workers in the agricultural sector \( (N_a) \), the average wage \( (\bar{W}) \), and the duration of formal and informal jobs \( (d_1 \text{ and } d_0, \text{ respectively}) \).

We can derive explicit expressions for those targets from our model, and the data values for those targets are readily observed from the household survey. Using these target values, we compute the values of \( \lambda, \delta, \alpha, m(\theta) \) and \( A_1 \). Next, we determine the migration cost \( (M) \) and the cost of a vacancy \( (c) \) using the no-migration and free-entry conditions, respectively. In what follows, we describe this calibration step in greater detail. We first show the analytic expressions for the targets.

Our model implies that informal sector workers face the job destruction rate \( \delta \). Because we are assuming that the rate follows a Poisson process, each employed worker in the informal sector is expected to hold his job for \( 1/\delta \) periods. In other words, the duration of a bad job is given by \( d_0 = 1/\delta \). We can calculate the duration of informal jobs from the household survey, and this allows us to back out \( \delta \).

The share of the population employed in the informal sector is given by
\[ n_0 = \int_{y^*}^{1} n_0(y) f(y) dy, \]  
(2)

with \( n_0(y) \) taking different values depending on the type of worker. In other words, for \( y < y^* \),

\[ n_0(y) = \frac{\alpha}{\delta + \alpha}, \]

for \( y^* < y < y^{**} \),

\[ n_0(y) = \frac{\alpha \lambda G(R(y))}{\lambda (\delta + \alpha) G(R(y)) + \delta m(\theta) G(y)}, \]

and for \( y > y^{**} \),

\[ n_0(y) = 0. \]

The share of the population employed in the formal sector is given by

\[ n_1 = \int_{y^*}^{1} n_1(y) f(y) dy, \]  
(3)

with \( n_1(y) \) taking different values depending on the worker type. For \( y < y^* \),

\[ n_1(y) = 0, \]

for \( y^* < y < y^{**} \),

\[ n_1(y) = \frac{\delta m(\theta) G(y)}{\lambda (\delta + \alpha) G(R(y)) + \delta m(\theta) G(y)}, \]

and for \( y > y^{**} \),

\[ n_1(y) = \frac{m(\theta) G(y)}{\lambda G(R(y)) + m(\theta) G(y)}. \]
The share of workers in agriculture is given by

\[ N_a = F(y^r). \] (4)

The parameters of the skill distribution are obtained as described in Step 1, and because we observe the share of workers in agriculture directly from the survey, we can back out \( N_a \).

The duration of employment in the formal sector \( (d_1) \) is given by

\[ d_1 = \int_{\max[y^*, y^r]}^{\max[y^*, y^r]} \frac{1}{h_1(y)} \frac{n_1(y)f(y)}{n_1} \, dy, \] (5)

where \( h_1(y) = \lambda \frac{G(R(y))}{G(y)} \) is the unemployment hazard rate for workers of type \( y \) and \( \frac{n_1(y)f(y)}{n_1} \) is the density of types among employed workers in the formal sector. This expression comes from the fact that for a worker employed in the formal sector, a shock arrives at the rate of \( \lambda \). However, the job is destroyed only if new productivity falls below the reservation productivity \( R(y) \). This means that every period, a worker of type \( y \) faces a probability \( h_1(y) = \lambda \frac{G(R(y))}{G(y)} \) of losing his or her job. This is known as the hazard rate or exit rate out of employment. Because we are assuming a constant shock arrival rate, the duration of employment in the formal sector for a worker of type \( y \) is the inverse of his/her hazard rate \( (d_1(y) = 1/h_1(y)) \). The lower integrand takes into account that no worker of type \( y < y^* \) would hold a formal job and that only workers of type \( y > y^r \) would be in the urban economy.

Finally, the average wage is given by\(^{27}\):

\(^{27}\)See ANV 2009 for the derivation of this expression
\[
W = \int_{y^*}^{1} \Omega(y) \frac{n_1(y)f(y)}{n_1} dy
\]

\[
\Omega(y) = w(y,y) \frac{G(R(y))}{G(y)} + \int_{R(y)}^{y} w(x,y) \frac{g(x)}{G(y)} dx.
\]

We use the resulting value \(m(\theta)\) together with the predefined values of \(\rho = 1/2\) and \(\zeta = 1\) to find the value of \(\theta\) by \(m(\theta) = \zeta \theta^\rho\). We now can use the free entry condition (1.22) to back out the cost of posting a vacancy \(c\). Once we know \(\theta\), we can use the expression for unemployment to find the implied vacancy rate in the economy.
Appendix C: Computation Procedure in Chapter 3

We solve the dynamic programming problem using a collocation method. Because our value function is continuous on a compact, convex interval, we can approximate the value function with polynomials. We use Chebychev polynomials for our basis functions.\(^{28}\)

First we convert AR(1) learning ability into a discrete Markov process using Tauchen’s method. We choose five discrete points and use the transition matrix to indicate conditional probability weights. Next, we approximate the value function using \(n\) basis functions at \(n\) collocation nodes.

We choose 100 nodes and thus the 99th degree of Chebychev polynomial function. Because the polynomial approximant behaves strangely at extrapolated points, we must choose an adequately large state interval. We choose the interval \([1, 1000]\) for our benchmark model. The maximum approximation error for our benchmark specification is about \(0.4 \times 10^{-3}\) and becomes less than \(10^{-4}\) for a wide range of the specified interval.

The computational procedure is as follows.

**Step 1** Guess a set of parameters \(\{\gamma, \rho, \sigma_v\}\).

**Step 2** Approximate the policy function for each level of learning ability

**Step 3** Guess \(F^0(\log(y), \log(a))\)

**Step 4** If \(F^0\) converges to the stationary distribution \(\bar{F}\), the iteration stop. Otherwise, repeat the process, starting with Step 1.

\(^{28}\text{It is well known that Chebychev-node polynomial interpolants are relatively accurate given their low approximation errors (see Rivlin’s theorem), and even with high-order polynomial approximants, Chebychev polynomials are quite accurate and computationally efficient (Fackler, 2002).}\)
We fix $\beta = (\frac{1}{1+0.04})^{20}$. Other parameters in step 1 are determined within the model as described in section 4. In step 2, we approximate the value function using the Chevyshev polynomial, whereas the policy function is computed using the policy function iteration algorithm. For this computational procedure, we use the dynamic programming toolbox developed by Miranda and Fackler (2002).

In Step 3, we first generate the learning ability paths for 1000 simulation periods using the stationary log($a$) distribution and the AR(1) process. We convert the continuous values to discrete ones (i.e., a discrete approximation of the ability paths). Next, we choose an arbitrary log($y$) distribution with a large number of observations. By design, the distribution of learning ability is stationary. Given the policy function computed in step 2, we generate a large number of observations $\{(y_i, a_i)\}_{i=1}^N$, and we generate a simulation using the policy function spline while computing $\text{var}(\log(y))$, $Gini(\log(y))$, and $\text{corr}(\log(y), \log(y'))$. If those moments converge to the data moments, the process is complete. Otherwise, we repeat the process, starting from step 1.
Table 1: Distribution of Workers Across Labor Market States

<table>
<thead>
<tr>
<th>Sector</th>
<th>Share of the Labor Force (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unemployed</td>
<td>11.29</td>
</tr>
<tr>
<td>Informal Sector</td>
<td>21.74</td>
</tr>
<tr>
<td>Formal Sector</td>
<td>36.30</td>
</tr>
<tr>
<td>Agriculture</td>
<td>30.67</td>
</tr>
</tbody>
</table>

Table 2: Target Statistics

<table>
<thead>
<tr>
<th>Target Statistics</th>
<th>Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Share of Formal Sector Employment</td>
<td>36.30%</td>
</tr>
<tr>
<td>Share of Informal Sector Employment</td>
<td>21.74%</td>
</tr>
<tr>
<td>Duration of a Formal Sector Job</td>
<td>6.17 years</td>
</tr>
<tr>
<td>Average Formal Sector Wage</td>
<td>1.1618</td>
</tr>
</tbody>
</table>
Table 3: Calibrated Parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost of Posting a Vacancy</td>
<td>0.4163</td>
</tr>
<tr>
<td>Migration Cost</td>
<td>0.3745</td>
</tr>
<tr>
<td>Arrival Rate of Informal Sector Employment</td>
<td>0.4895</td>
</tr>
<tr>
<td>Shock Arrival Rate in Informal Sector</td>
<td>0.176</td>
</tr>
<tr>
<td>Shock Arrival Rate in Formal Sector</td>
<td>0.2342</td>
</tr>
<tr>
<td>Output Elasticity of Employment in Agriculture</td>
<td>0.628</td>
</tr>
<tr>
<td>Fixed Income from Informal Sector</td>
<td>0.9767</td>
</tr>
<tr>
<td>Technological Parameter of Agricultural Function</td>
<td>0.4966</td>
</tr>
<tr>
<td>Technological Parameter of Formal Sector Function</td>
<td>2.8812</td>
</tr>
<tr>
<td>Parameter of Skill Distribution</td>
<td>0.9621</td>
</tr>
<tr>
<td>Parameter of Skill Distribution</td>
<td>0.7914</td>
</tr>
<tr>
<td>Other Parameter Value</td>
<td></td>
</tr>
<tr>
<td>Other Parameter</td>
<td>Value</td>
</tr>
<tr>
<td>-----------------------------------------------------</td>
<td>--------</td>
</tr>
<tr>
<td>Average Income in Agriculture</td>
<td>$y_a$</td>
</tr>
<tr>
<td>Average Income in Formal Sector (average wage)</td>
<td>$y_1 = W$</td>
</tr>
<tr>
<td>Unemployment Rate</td>
<td>$u$</td>
</tr>
<tr>
<td>Threshold Level of Skills required to migrate</td>
<td>$y^f$</td>
</tr>
<tr>
<td>Threshold Level of Skills for Informal Sector Search</td>
<td>$y^*$</td>
</tr>
<tr>
<td>Threshold Level of Skills for Formal Sector Search</td>
<td>$y^{**}$</td>
</tr>
</tbody>
</table>

Table 4: Benchmark Values of Endogenous Variables
<table>
<thead>
<tr>
<th>Simulation</th>
<th>% dp by ( \lambda ) up by ( 5% )</th>
<th>% dp by ( y_0 ) up by ( 5% )</th>
<th>% dp by ( \delta ) up by ( 5% )</th>
<th>% dp by both ( A_1 ) and ( y_0 ) down by ( 5% )</th>
<th>% dp by both ( A_1 ) and ( y_0 ) down by ( 10% )</th>
<th>% dp by both ( A_1 ) and ( y_0 ) up by ( 5% ) and ( \lambda ) up by ( 5% )</th>
<th>% dp by both ( A_1 ) and ( y_0 ) up by ( 5% ) and ( \delta ) up by ( 5% )</th>
<th>% dp by both ( A_1 ) and ( y_0 ) up by ( 5% ) and ( \lambda ) up by ( 5% ) and ( \delta ) up by ( 5% )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benchmark</td>
<td>0.3647</td>
<td>0.2164</td>
<td>0.3069</td>
<td>0.112</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Simulation 1</td>
<td>0.3602</td>
<td>0.2202</td>
<td>0.3069</td>
<td>0.1127</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Simulation 2</td>
<td>0.3553</td>
<td>0.2244</td>
<td>0.3069</td>
<td>0.1134</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Simulation 3</td>
<td>0.29</td>
<td>0.321</td>
<td>0.2564</td>
<td>0.1326</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Simulation 4</td>
<td>0.3567</td>
<td>0.223</td>
<td>0.3069</td>
<td>0.1134</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Simulation 5</td>
<td>0.3537</td>
<td>0.2254</td>
<td>0.3069</td>
<td>0.1139</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Simulation 6</td>
<td>0.3695</td>
<td>0.2102</td>
<td>0.3069</td>
<td>0.1134</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Simulation 7</td>
<td>0.3743</td>
<td>0.2041</td>
<td>0.3069</td>
<td>0.1147</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 5: Simulation: Reallocation of Workers Across Sectors
<table>
<thead>
<tr>
<th>Benchmark</th>
<th>Decrease in Total Employment (%)</th>
<th>Change in Per Capita Income (%)</th>
<th>Change in Average Formal Sector Wage (%)</th>
<th>Duration of a Formal Sector Job (months)</th>
<th>Duration of Unemployment (months)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simulation 1</td>
<td>Both $A_1$ and $y_0$ down by 5%</td>
<td>0.079</td>
<td>-4.65</td>
<td>77.93</td>
<td>32.05</td>
</tr>
<tr>
<td>Simulation 1</td>
<td>Both $A_1$ and $y_0$ down by 10%</td>
<td>0.16</td>
<td>-9.28</td>
<td>79.90</td>
<td>33.01</td>
</tr>
<tr>
<td>Simulation 1</td>
<td>$A_1$ down by 10% &amp; $y_0$ up by 5%</td>
<td>2.32</td>
<td>-3.41</td>
<td>84.49</td>
<td>36.17</td>
</tr>
<tr>
<td>Simulation 2</td>
<td>$\lambda$ up by 5%</td>
<td>0.16</td>
<td>-0.98</td>
<td>75.45</td>
<td>31.92</td>
</tr>
<tr>
<td>Simulation 2</td>
<td>$\lambda$ up by 10%</td>
<td>0.21</td>
<td>-1.31</td>
<td>73.08</td>
<td>31.72</td>
</tr>
<tr>
<td>Simulation 3</td>
<td>$\delta$ up by 5%</td>
<td>0.16</td>
<td>-0.31</td>
<td>76.35</td>
<td>31.33</td>
</tr>
<tr>
<td>Simulation 3</td>
<td>$\delta$ up by 10%</td>
<td>0.30</td>
<td>-0.60</td>
<td>76.47</td>
<td>31.45</td>
</tr>
<tr>
<td>Simulation 4</td>
<td>Both $A_1$ and $y_0$ down by 10% &amp; $\lambda$ up by 5%</td>
<td>0.23</td>
<td>-9.61</td>
<td>77.35</td>
<td>32.82</td>
</tr>
<tr>
<td>Simulation 4</td>
<td>$A_1$ down by 10% &amp; $y_0$ up by 5% &amp; $\lambda$ up by 5%</td>
<td>2.44</td>
<td>-4.20</td>
<td>84.96</td>
<td>36.93</td>
</tr>
<tr>
<td>Simulation 5</td>
<td>Both $A_1$ and $y_0$ down by 10% &amp; $\lambda$ up by 5% &amp; $\delta$ up by 5%</td>
<td>0.39</td>
<td>-9.91</td>
<td>77.49</td>
<td>32.97</td>
</tr>
<tr>
<td>Simulation 5</td>
<td>$A_1$ down by 10% &amp; $y_0$ up by 5% &amp; $\lambda$ up by 5% &amp; $\delta$ up by 5%</td>
<td>2.38</td>
<td>-5.05</td>
<td>83.92</td>
<td>36.93</td>
</tr>
</tbody>
</table>

Table 6: Simulation: Changes in Labor Markets Variables
Table 7: Unemployment Rate

<table>
<thead>
<tr>
<th>YEAR</th>
<th>IFLS (%)</th>
<th>Sakernas (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1997</td>
<td>7.60</td>
<td>4.70</td>
</tr>
<tr>
<td>2000</td>
<td>2.02</td>
<td>6.3</td>
</tr>
</tbody>
</table>

Table 8: Employment Composition and Unemployment Rate

<table>
<thead>
<tr>
<th></th>
<th>1997</th>
</tr>
</thead>
<tbody>
<tr>
<td>Share of 'active' working age population employed in agriculture</td>
<td>33.16</td>
</tr>
<tr>
<td>Share of the employed in informal sector</td>
<td>26.25</td>
</tr>
<tr>
<td>Share of the employed in formal sector</td>
<td>35.89</td>
</tr>
<tr>
<td>Share of 'active' working age unemployed population (narrow)</td>
<td>4.70</td>
</tr>
</tbody>
</table>

Table 9: Labor Income across Sectors

<table>
<thead>
<tr>
<th></th>
<th>1997</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average Income in Agriculture (IDR per month)</td>
<td>368,576</td>
</tr>
<tr>
<td>Average Income in Informal Employment (IDR per month)</td>
<td>751,117</td>
</tr>
<tr>
<td>Average Income in Formal Employment (IDR per month)</td>
<td>1,062,819</td>
</tr>
</tbody>
</table>
### Table 10: Benchmark Parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Notation</th>
<th>Value</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Discount Rate</td>
<td>r</td>
<td>0.1381</td>
<td>IMF</td>
</tr>
<tr>
<td>a parameter for Skill in Beta Distribution</td>
<td>$\beta_a$</td>
<td>1.6593</td>
<td>IFLS</td>
</tr>
<tr>
<td>b parameter for Skill in Beta Distribution</td>
<td>$\beta_b$</td>
<td>1.8106</td>
<td>IFLS</td>
</tr>
<tr>
<td>Leisure Value of Unemployment</td>
<td>b</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Average Informal Labor Income</td>
<td>$y_0$</td>
<td>0.7511</td>
<td>IFLS</td>
</tr>
<tr>
<td>Output Elasticity of Labor in Agriculture</td>
<td>$\gamma$</td>
<td>0.51</td>
<td>Bravo-Ortega and Lederman (2004)</td>
</tr>
<tr>
<td>Agricultural TFP</td>
<td>$A_a$</td>
<td>0.2146</td>
<td>Model</td>
</tr>
<tr>
<td>Formal Sector TFP</td>
<td>$A_1$</td>
<td>1.8735</td>
<td>Model</td>
</tr>
<tr>
<td>Shock Arrival Rate of Formal Jobs</td>
<td>$\lambda$</td>
<td>0.3021</td>
<td>Model</td>
</tr>
<tr>
<td>Informal Job Arrival Rate</td>
<td>$\alpha$</td>
<td>4.0163</td>
<td>Model</td>
</tr>
<tr>
<td>Migration Cost</td>
<td>$M$</td>
<td>0.3281</td>
<td>Model</td>
</tr>
<tr>
<td>Cost of formal Vacancy</td>
<td>c</td>
<td>0.3370</td>
<td>Model</td>
</tr>
<tr>
<td>Informal Job Destruction Rate</td>
<td>$\delta$</td>
<td>0.1757</td>
<td>GPP (2008)</td>
</tr>
<tr>
<td>Matching Efficiency Parameter</td>
<td>$\zeta$</td>
<td>1</td>
<td>Standard</td>
</tr>
</tbody>
</table>

### Table 11: Benchmark Outcomes

<table>
<thead>
<tr>
<th>Variable</th>
<th>Notation</th>
<th>Data</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unemployment Rate</td>
<td>$U$</td>
<td>0.047</td>
<td>0.047</td>
</tr>
<tr>
<td>Share of Workers in Agriculture</td>
<td>$N_a$</td>
<td>0.3316</td>
<td>0.3316</td>
</tr>
<tr>
<td>Share of Workers in Informal Sector</td>
<td>$N_0$</td>
<td>0.2625</td>
<td>0.2625</td>
</tr>
<tr>
<td>Share of Workers in Formal Sector</td>
<td>$N_1$</td>
<td>0.3589</td>
<td>0.3589</td>
</tr>
<tr>
<td>Average Income in Agriculture</td>
<td>$Y_a$</td>
<td>0.3686</td>
<td>0.3686</td>
</tr>
<tr>
<td>Average Wage in Formal Sector</td>
<td>$w$</td>
<td>1.0628</td>
<td>1.0628</td>
</tr>
<tr>
<td>Average Income in the Economy</td>
<td>$Y$</td>
<td>0.7008</td>
<td></td>
</tr>
<tr>
<td>Duration of Unemployment</td>
<td>$d(u)$</td>
<td>0.8069</td>
<td>0.8069</td>
</tr>
<tr>
<td>Market Tightness</td>
<td>$\theta$</td>
<td>2.3349</td>
<td></td>
</tr>
<tr>
<td>Vacancy Rate</td>
<td>$\nu$</td>
<td>0.1097</td>
<td></td>
</tr>
<tr>
<td>No-migration Threshold Value</td>
<td>$y_r$</td>
<td>0.3514</td>
<td></td>
</tr>
<tr>
<td>Cut-off Value of Formal Job Taking</td>
<td>$y^*$</td>
<td>0.3719</td>
<td></td>
</tr>
<tr>
<td>Cut-off Value of Informal Job Taking</td>
<td>$y^{**}$</td>
<td>0.6015</td>
<td></td>
</tr>
<tr>
<td>Formal Job Matching Rate</td>
<td>$m(\theta)$</td>
<td>1.5281</td>
<td></td>
</tr>
</tbody>
</table>
Table 12: Avg. Education by the Model and Data Observation

<table>
<thead>
<tr>
<th>Data</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Avg. Education in Agricultural Sector</td>
<td>5.44 3.3768</td>
</tr>
<tr>
<td>Avg. Education in Informal Employed Workers</td>
<td>6.61 7.5417</td>
</tr>
<tr>
<td>Avg. Education for Formal Employed Workers</td>
<td>9.01 11.2506</td>
</tr>
<tr>
<td>Avg. Education for Unemployed Workers</td>
<td>8.68 10.9311</td>
</tr>
</tbody>
</table>

Table 13: Benchmark Parameters

<table>
<thead>
<tr>
<th>β</th>
<th>γ</th>
<th>ρ</th>
<th>σv</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.4564</td>
<td>0.475</td>
<td>0.1</td>
<td>0.5</td>
</tr>
</tbody>
</table>

Table 14: Target and Model Moments in Benchmark Case

<table>
<thead>
<tr>
<th>Moment</th>
<th>Target</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>corr(log(y'), log(y))</td>
<td>0.5</td>
<td>0.51</td>
</tr>
<tr>
<td>var(log(y))</td>
<td>0.36</td>
<td>0.356</td>
</tr>
<tr>
<td>Gini(log(y))</td>
<td>0.301</td>
<td>0.31</td>
</tr>
</tbody>
</table>
Table 15: International Comparison

<table>
<thead>
<tr>
<th>Country</th>
<th>$\beta_{\text{data}}$</th>
<th>$Gini_{\text{data}}$</th>
<th>$\beta_{\text{model}}$</th>
<th>$Gini_{\text{model}}$</th>
<th>$\gamma$</th>
<th>$\sigma_v$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Australia</td>
<td>.314</td>
<td>.281</td>
<td>.314</td>
<td>.280</td>
<td>.247</td>
<td>.483</td>
</tr>
<tr>
<td>Canada</td>
<td>.2</td>
<td>.281</td>
<td>.2029</td>
<td>.280</td>
<td>.125</td>
<td>.5</td>
</tr>
<tr>
<td>Denmark</td>
<td>.123</td>
<td>.254</td>
<td>.123</td>
<td>.253</td>
<td>.041</td>
<td>.473</td>
</tr>
<tr>
<td>Finland</td>
<td>.277</td>
<td>.209</td>
<td>.278</td>
<td>.209</td>
<td>.194</td>
<td>.388</td>
</tr>
<tr>
<td>France</td>
<td>.41</td>
<td>.293</td>
<td>.408</td>
<td>.294</td>
<td>.347</td>
<td>.485</td>
</tr>
<tr>
<td>Germany</td>
<td>.326</td>
<td>.244</td>
<td>.326</td>
<td>.245</td>
<td>.25</td>
<td>.433</td>
</tr>
<tr>
<td>Italy</td>
<td>.435</td>
<td>.306</td>
<td>.435</td>
<td>.306</td>
<td>.38</td>
<td>.496</td>
</tr>
<tr>
<td>Norway</td>
<td>.246</td>
<td>.223</td>
<td>.246</td>
<td>.223</td>
<td>.163</td>
<td>.412</td>
</tr>
<tr>
<td>Sweden</td>
<td>.24</td>
<td>.197</td>
<td>.244</td>
<td>.195</td>
<td>.16</td>
<td>.372</td>
</tr>
<tr>
<td>UK</td>
<td>.5</td>
<td>.27</td>
<td>.510</td>
<td>.276</td>
<td>.454</td>
<td>.44</td>
</tr>
</tbody>
</table>

Note: For the estimates of income correlation (elasticity) and cross-sectional income inequality (Gini coefficients), we used Table Y in Björklund and Jäntti (2009) where they survey the cross-country income correlations (elasticities). For the Gini coefficient, they refer to LIS Wave 1 or 2 (around 1980 if available, around 1985 if the wage 1 data are not available).
Figure 1: Labor Market Structure
Figure 2: Case 1: No-Migration
Figure 3: Case 2: No-Migration
Figure 4: Case 1: Equilibrium with Rural-to-Urban Migration

\[ V(yr, \Theta) = 0 \]

\[ Na(y) = U(y) - M \]
Figure 5: Case 2: Equilibrium with Rural-to-Urban Migration
Figure 6: No-migration
Figure 7: Reservation Productivity
Figure 8: Skill Distribution
Figure 9: Formal Wage Distribution
Figure 10: Employment Share
Figure 11: Skill Distribution across Sectors
Figure 12: Unemployment Share Distribution across Skill
Figure 13: Cross-Country Gini coefficient vs. $\sigma_v$
Figure 14: Intergenerational log-income correlation vs $\gamma$
Figure 15: Intergenerational Log-income Correlation VS Government Educational Spending Ratio To GDP
Figure 16: Intergenerational Log-Income Correlation with $\sigma_v$
Figure 17: Intergenerational Log-Income Correlation with Direct Government Human Capital Investment Subsidy
Figure 18: Approximation Error