Spring 5-15-2020

Essays on Growth, Development, and Human Capital

Juan Ignacio Vizcaino
Washington University in St. Louis

WASHINGTON UNIVERSITY IN ST. LOUIS

Department of Economics

Dissertation Examination Committee:
Rodolfo Manuelli (Chair)
Francisco Buera
Alexander Monge-Naranjo
Yongseok Shin
Ping Wang

Essays on Growth, Development, and Human Capital

by

Juan Ignacio Vizcaino

A dissertation presented to
The Graduate School
of Washington University in
partial fulfillment of the
requirements for the degree
of Doctor of Philosophy

May, 2020
St. Louis, Missouri

Table of Contents

List of Figures

List of Tables

Acknowledgements

I am deeply indebted to Rodolfo Manuelli for his patience and continuous support during my years as a PhD student. His guidance and insights have undoubtedly influenced the present dissertation. I also thank Francisco Buera for various discussions on the main chapter and close collaboration on the second chapter of the dissertation. This dissertation was further enriched by comments and suggestions by Ping Wang, Yongseok Shin, and Limor Golan whom I thank for their involvement too. Finally, I thank the Graduate School of Washington University for the financial support.

Juan Ignacio Vizcaino

Washington University in St. Louis

May, 2020

ABSTRACT OF THE DISSERTATION

Essays on Growth, Development, and Human Capital

by

Juan Ignacio Vizcaino

Doctor of Philosophy in Economics

Washington University in St. Louis, 2020

Professor Rodolfo Manuelli, Chair

Skills, Technologies and Development. I study how the productivity of skilled and unskilled labor varies with development. Using harmonized, occupational labor market outcomes for a broad set of countries across the development spectrum, I document that employment in high-skill occupations, or jobs that are relatively more intensive in non-routine cognitive tasks, grows with development. In addition, the income of workers in high-skill occupations falls relative to earnings in low-skill occupations as countries grow richer. To understand the forces driving these findings, I develop a stylized model of the labor market across development. In the model, labor productivity is determined endogenously as a result of the selection of heterogeneous workers into occupations and education. I use a quantitative version of the model to decompose the observed decline in relative labor income between less-developed countries and the US into a component embedded in technologies, or relative skilled labor efficiency, and a fraction due to workers' characteristics, or relative skilled labor quality. I find that relative quality explains 25 percent of the decline in relative labor income, with the remaining fraction due to relative efficiency. In less-developed

countries, the relatively few skilled workers are the most productive in performing high-skill jobs, which reduces the magnitude of skill-biased technological progress needed to rationalize the cross-country data by one half when compared to a world where labor quality is purely determined by educational attainment.

Skill-Biased Structural Change. Using a broad panel of advanced economies, we document that increases in GDP per-capita are associated with a systematic shift in the composition of value added to sectors that are intensive in high-skill labor, a process we label as skill-biased structural change. It follows that further development in these economies leads to an increase in the relative demand for skilled labor. We develop a two-sector model of this process and use it to assess the contribution of skill-biased structural change to the rise of the skill premium in the US and a set of ten other advanced economies, over the period of 1977 to 2005. For the US, we find that these compositional changes in demand account for 20-27% of the overall increase of the skill premium due to technical change.

Natural Disasters and Growth: The Role of Foreign Aid and Disaster Insurance. In this paper we develop a continuous time stochastic growth model that is suitable for studying the impact of natural disasters on the short run and long run growth rate of an economy. We find that the growth effects of a natural disaster depend in complicated ways on the details of expected foreign disaster aid and the existence of catastrophe insurance markets. We show that aid can have an influence on investments in prevention and mitigation activities and can delay the recovery from a natural disaster strike.

Chapter 1

Skills, Technologies, and Development

Juan Ignacio Vizcaino

1.1 Introduction

Cross-country differences in standards of living and labor force attributes are noticeable. Understanding what shapes the observed disparities in workers' qualifications and how they translate into different prosperity paths is a central element in the analysis of economic development. The traditional view is that there are sizable gaps in factor-neutral productivity across countries. More recent studies, based on newly available data and improved measuring techniques find that technological progress is biased towards skilled labor, reflecting a shift in demand in favor of workers with higher levels of educational attainment along the development path (Caselli and Coleman (2006), Jones (November 2014), Caselli (2005), Rossi (2017), Malmberg (2018)). This relative demand shift is needed in order to reconcile the cross-country empirical fact that large improvements in educational attainment are associated with a relatively small decline in the educational skill premium as countries develop.

Strikingly, measured skill bias in cross-country technological efficiency is of such magnitude that models where workers' labor productivity is solely determined by their educational attainment predict that higher-income countries use unskilled labor not only relatively but also absolutely less efficiently.

In this paper I revisit this issue and study how the relative efficiency of the technologies that use skilled and unskilled labor and the quality of skilled and unskilled labor vary with development. Consistent with new data on occupational labor income across the development spectrum, I find that allowing for heterogeneity in workers' abilities reduces the magnitude of measured skill-bias technological progress by one half when compared to a world where differences in the quality of the labor force are purely determined by educational attainment. As a consequence, unlike the most-recent strand of the literature, more-developed countries use both skilled and unskilled labor more efficiently.

The key difference between my work and previous studies is in the modeling and measurement of the relative quality of skilled and unskilled labor. The literature typically assumes that differences in the quality of workers are mostly explained by their educational attainment, which implicitly considers that individuals are homogeneous in their characteristics once differentials in schooling levels are accounted for. In my case, instead, I allow for workers to have heterogeneous attributes, which together with the state of technology shape their occupational and educational choices and determine the quality of skilled and unskilled labor in equilibrium.

My paper also introduces a novel empirical approach. Earlier studies often rely on cross-country data on educational attainment and Mincerian returns to education to disentangle the behavior of relative skilled labor productivity and efficiency along the development path.

Thus, splitting countries' labor forces into skilled and unskilled employees requires choosing a minimum level of education that differentiates workers qualitatively into different production factors. In my case, I group workers according to the main occupation they perform, which more closely reflects the task component of their jobs and more clearly reflects qualitative differences between them.

Following this alternative classification criterion, I use harmonized, occupational labor market outcomes for a broad set of countries across the development spectrum and document that employment in high-skill occupations, or jobs that are relatively more intensive in non-routine cognitive tasks, grows with development. In addition, workers earnings in high-skill occupations falls with respect to those in low-skill occupations as countries grow richer, with elasticities in line with those found by studies based on educational attainment and Mincerian returns to education.

To shed light on these findings and disentangle the mechanisms that determine the relative quality and efficiency of skilled labor, I build a general equilibrium model of occupational choice and human capital accumulation through education. The labor demand side is characterized by a representative, cost-minimizing firm that operates the single aggregate technology available in the economy. As in previous studies, the technology features a labor aggregator where skilled and unskilled labor are imperfect substitutes in production. Given this structure, exogenous skill-biased shifts in technological efficiency attract more workers towards high-skill occupations. A novel feature of the model is that the effective supply of skilled and unskilled labor is determined endogenously by workers' occupational choice and their decision to accumulate human capital through education. As a result, the equilibrium productivities of skilled and unskilled labor crucially depend on the properties of the joint distribution of skills in the population.

I perform a mixture of calibration and estimation of the model's deep parameters, including those governing the joint distribution of skills, to match some labor market moments in the US. I use the quantitative version of the model to conduct two exercises that highlight the importance of workers' attributes when making their occupational and educational decisions, and how they are related to skilled-biased technological change.

In the first exercise, I fix the parameters of the joint skill distribution and compute the levels of relative skilled labor efficiency and two model parameters that capture fixed costs of schooling that are required to rationalize the observed shares of skilled workers and educational attainment across countries. As a by-product, I obtain the non-targeted levels of relative labor income in high- and low-skill occupations. Qualitatively, the model reproduces the decline and the pattern of relative labor income in high- and low-skill occupations we observe in the data as countries develop. Quantitatively, my framework explains 70 percent of the total decline, which I decompose into a relative efficiency and a relative quality component. I find that between 25 percent of the observed relative labor income differentials are explained by the relative quality of skilled labor, while the remaining fraction is due to skill-biased technological progress.

A natural outcome of this exercise is to compare the model predictions for technological efficiency gaps between countries. When I do so, I find a sizeable skill-bias in technological progress as countries develop, with relative labor efficiency being more than a hundred times higher in the US than in the set of least developed countries in my sample, and about fifty times higher than in the average country in the second development quartile. Since my sample encompasses a broader set of countries at the lower end of the development spectrum, the latter group is more relevant for comparison with earlier studies. The fact that the relative quality of skilled labor is higher in poor countries dampens the measured differences in skill-bias between rich and poor countries when compared to a model that measures labor quality based exclusively on educational attainment. In a world where labor quality is purely determined by educational attainment, the measured gap in relative efficiency between rich and poor countries would be two times larger, or 120 times higher in the US than in the average country in the second development quartile.

In the second main quantitative exercise, I investigate the role of an educational expansion. To asses the effects of increased access to education on development outcomes, I perform a reduction in the costs of acquiring education for the average country in the first development quartile. The engineered expansion is such that, after the thought policy is in place, the educational attainment levels in least-developed countries are the same as those observed for the average country in the most-developed group. In line with the results of a major educational expansion that took place in Brazil between 1995 and 2014 (see Jaume (2019)), I find that the occupational structure remains fairly unchanged compared to the educational attainment structure, with workers of all educational groups increasingly employed in lower-wage occupations.

In addition, the model predicts a growth in output per worker in the order of ten percent, a relative small number in comparison with the observed income gaps between most- and least-developed nations. The effects on GDP per-worker are at least one-third smaller than what a model that classifies workers into high- and low-skill according to their educational attainment would anticipate. The reason is that, in my model, the expansion benefits workers in high-skill occupations proportionately more, which through general equilibrium effects reduces the fraction of workers in high-skill occupations after the policy is implemented, compensating the initial increase in the effective supply of high-skill labor. On the other hand, in the model based on educational attainment, the expansion leads to a direct increase in the effective supply of high-skill labor, without any reduction in quantities coming through the workers' selection channel.

The quantitative exercises all together highlight the importance the selection of workers into occupations and education together with skill-biased technological progress in shaping differences in standards of living across countries and improving labor market outcomes for both high- and low-skill workers.

Related Literature. This paper contributes to the strand of the economic development literature that studies skill-biased technological differences between countries. An early contribution is by Caselli and Coleman (2006), who propose an aggregate technology framework to unveil cross-country skilled-biased gaps in technological efficiency. Caselli (2016) updates and expands this study to a broader set of countries and other factors besides labor, but still based on an aggregate technology approach. More recently, Malmberg (2018) proposes a novel approach to estimate the relative efficiency of skilled and unskilled labor based on disaggregated trade and industry data. Another recent contribution is given by Rossi (2017), who compares labor market outcomes of immigrants with different levels of educational attainment to identify differences in the relative efficiency and the relative quality of skilled and unskilled labor. The main difference between my work and these studies is that I propose a model-based method to estimate the relative quality of skilled and unskilled labor. In addition, and except for the case of Malmberg, I use occupational attainment instead of educational attainment data to identify qualitative differences between workers.

My paper is also related to a large body of literature that finds evidence of skilled-biased technical change across time and within countries. Katz and Autor (1999) provide a comprehensive survey of this literature, that includes Katz and Murphy (1992a), Acemoglu (1998), Autor et al. (1998), Acemoglu (2002).

From a methodological perspective, a paper that is closely related to mine is Lagakos and Waugh (April 2013). However, in their case, the self-selection of workers of heterogeneous abilities into different sectors is used to explain differences agricultural productivities between rich and poor countries. Moreover, in their work countries only differ in an economy-wide efficiency parameter and there is no role for skill-biased technological progress.

From a broader point of view, this paper is also related to the literature that measures the contribution of human capital to development. Earlier contributors to this literature are Mankiw et al. (1992), Klenow and Rodriguez-Clare (1997), and Hall and Jones (1999). More recently, Erosa et al. (2010), Manuelli and Seshadri (September 2014), Jones (November 2014) find human capital to be an important factor in explaining disparities in wealth levels between countries. In my case, instead, I focus on productivity differentials between two groups of workers that are different in their nature due to the tasks they more commonly perform, rather than studying the role of labor quality as a whole.

The paper is organized as follows. Sections 1.2 and 1.3 present the empirical analysis, including sources, a detailed description of the data, and robustness checks. Section 1.4 presents the model. In order to build up some intuition, I show in Subsection 1.4.2 the model's basic structure, given by the workers' selection problem into occupations according to their *unobservable characteristics*. Subsection 1.4.5 shows the properties of the joint skill distribution that are key to understand what mechanisms in the model generate the patterns we observe in the data, while Subsection 1.4.6 builds endogenous human capital accumulation into the model. Section 1.5 describes the strategy followed to estimate the model parameters, while in Section 1.6.1 I present my main quantitative exercises. Section 1.7 concludes.

1.2 Empirical Analysis

1.2.1 Data Description

My main data source is the International Labor Organization (ILO) [1]. In particular, I use cross-country, harmonized, occupational level data on average nominal labor income of employees [2], average weekly hours worked per employee, and number of people employed.

The ILO provides occupational statistics at the one-digit level of aggregation following the International Standard Classification of Occupations (ISCO). I focus on countries that have data for the latest version of the ISCO classification, ISCO-08 [3]. ILO's aggregated statistics are based on micro-data that is representative of the labor force of each country. The micro-data comes from surveys or studies that vary between countries, depending on availability. Data sources include Labor Force Surveys, Employment Surveys, Establishment Surveys, Household Surveys, Insurance Records, and Administrative data. For countries with multiple data sources available, I prioritize labor force, employment and household surveys [4].

[1] See http://www.ilo.org/ilostat. Appendix A.1 presents a description of the ILO data, issues and treatment.

[2] More precisely, the ILO has data on labor earnings. The concept of earnings, as applied in wages statistics, relates to gross remuneration in cash and in kind paid to employees, as a rule at regular intervals, for time worked or work done together with remuneration for time not worked, such as annual vacation, other type of paid leave or holidays. Earnings exclude employers' contributions in respect of their employees paid to social security and pension schemes and also the benefits received by employees under these schemes.

[3] I discard data from ISCO-88 to avoid issues arising from methodological changes in occupational aggregation between the two ISCO releases. Even though the titles of the ten major occupational groups are the same under ISCO-08 and ISCO-88, some minor occupational groups were moved between major groups when ISCO-88 was updated to ISCO-08 to reflect the effects of technology on professional, technical, and clerical work. My sample is reduced in eleven observations by discarding countries with data available for ISCO-88 only.

[4] In these studies income is reported by workers rather than by employers, as in Administrative Data or Insurance Records. Thus, these type of studies provide a more accurate description of labor market outcomes

My main analysis takes into account workers of all ages [5] and both sexes [6]. Regarding the time frame, I focus on countries with information available between 2000 and 2018 and average data across time when information for multiple years is available for a preferred data source in any given country.

My cross-country empirical analysis and calibration uses data on GDP, capital, number of employees, and average hours worked per worker, which I obtain from Penn World Table 9.0 [7]. From this broad sample of GDP per worker for 182 countries I calculate the 25th, 50th, and 75th percentiles of GDP per worker, which I later on use to classify the 81 countries in my ILO sample into four development quartiles.

I exclude from my analysis countries intensive in natural resources, as defined in Appendix A.2, and those with a population smaller than one million. My sample has data for 81 countries and covers one-third of the world's population, including 20 countries in the first, 18 in the second, 23 in the third, and 20 in the fourth development quartiles. Regarding regional coverage, my sample includes 17 Advanced Economies, 10 from East Asia and Pacific, 10 from Europe and Central Asia, 15 from Latin America and the Caribbean, 4 from

for less-developed countries where the informal economy represents a large share of employment and it is more common for employees to receive payments under the table.

[5] It is customary in the empirical labor literature to restrict the analysis to either prime-aged workers or to workers with high labor force attachment. That is not possible in my case, as the ILO only provides summary statistics from for workers of all ages. My unit of analysis is employees instead of employed people, since the latter classification includes not only employees receiving remuneration but also working proprietors and unpaid family workers.

[6] It might be a concern that the inclusion of women might affect the analysis. For example, the level of attachment to the labor force could potentially differ for women who work in high- and low-skill occupations, especially at low development levels. Another concern could arise from the a gender wage gap that varies across development, especially in contexts of weak labor institutions. To tackle these type of issues I perform a robustness analysis in Section 1.3.2.

[7] The variables used are Output-side real GDP at chained PPPs in millions of 2011 US$ *(rgdpo)*, people engaged *(emp)*, average annual hours worked by persons engaged *(avh)*, and capital stock at 2011 national prices in millions of 2011 US$ *(rnna)*. See Feenstra et al..

Middle East and North Africa, 4 from South Asia, and 15 from Sub-Saharan Africa.

1.2.2 Occupational Aggregation.

The ILO provides labor market data by occupation at the one-digit level, ISCO-08's highest degree of aggregation. At that level occupations are collected into ten major groups [8]. I discard workers in Armed Forces Occupations, since their labor market outcomes might not necessarily be determined by market forces.

To simplify the analysis, I put together the nine major occupational groups left into two broader categories, which I label as high-skill and low-skill occupations. Behind this aggregation procedure lies the assumption that occupations are qualitatively different between groups. One way of capturing differences in nature between groups is to separate jobs according to the tasks that are more commonly performed in them.

Acemoglu and Autor (2011a) provide a clear characterization of the task intensity performed by each occupation. While professional, managerial and technical jobs are relatively more intensive in abstract, non-routine cognitive tasks, clerical and sales occupations, production workers and operators, and service workers more commonly perform routine cognitive, routine manual, and non-routine manual tasks. In turn, they document that for the US, job abstract intensity is positively and monotonically related to the skill level of an occupation, as approximated by the average wage of the workers in them. In addition, workers with higher educational attainment levels tend to concentrate proportionately more in occupations that are abstract task intensive.

[8] These major categories are: 1. Managers, 2. Professionals, 3. Technicians and Associate Professionals, 4. Clerical Support Workers, 5. Service and Sales Workers, 6. Skilled Agricultural, Forestry and Fishery Workers, 7. Craft and Related Trades Workers, 8. Plant and Machine Operators, and Assemblers, 9. Elementary Occupations, and 10. Armed Forces Occupations.

Non-routine cognitive task intensity of occupations seems to provide a good criterion to group them into high- and low-skill groups. As an additional robustness check to this criterion, I compare occupational wages at different development stages to see if high-skill occupations are indeed the ones with higher wages, as Acemoglu and Autor document to be the case for the US. To do so, I use ILO data to compute hourly labor income at purchasing power parity by country and occupation. I group the countries in my sample into four development quartiles, using the GDP per capita thresholds described in Subsection 1.2.1 above, and calculate the median hourly labor income at PPP for each quartile.

Table 1.1: Occupational Skill-Intensity
(median hourly labor income at PPP in US$ of 2011 by development quartile)

Broad Group	**High Skill**			**Low Skill**					
ISCO Occupation Title	*1. Managers*	*2. Professionals*	*3. Technicians*	*4. Clerical*	*5. Machine Operators*	*6. Craft Workers*	*7. Service Workers*	*8. Skilled Agricultural Workers*	*9. Elementary Occupations*
Bottom Quartile	4.8	3.8	3.0	2.7	1.9	1.9	2.7	2.3	1.7
Second Quartile	7.2	6.1	5.0	3.7	2.5	2.4	2.7	3.0	2.2
Third Quartile	21.4	14.5	10.0	6.9	4.8	4.3	5.8	5.9	4.1
Top Quartile	39.1	32.3	25.3	22.2	18.2	15.8	20.6	19.9	18.7

As Table 1.1 above shows, managers, professionals and technicians are indeed the highest earning occupations at all development levels. Thus, based both on abstract task intensity and the fact that they exhibit a higher wage level at all development quartiles I choose managers, professionals and technicians to be the occupations in my high-skill group [9].

1.2.3 Skill-Premium Definition.

I here explain how I construct my skill-premium measure, which I call occupational skill-premium. I start by computing hourly labor income for each of the nine occupational categories by diving monthly labor income, expressed at local currency units, by average monthly

[9] To be precise, high-skill occupations include ISCO-08's major groups 1,2, and 3. low-skill occupations are those in groups 4,5,6,7,8, and 9.

hours worked. The latter is obtained using data on average weekly hours worked and assuming that individuals work four weeks per month.

I define my skill-premium measure to be the ratio of the employment-weighted average hourly labor income for the ISCO categories in high-skill occupations with respect to the corresponding employment-weighted average labor income for the occupations in low-skill occupations.

I consider this to be an improvement with respect to previous work for several reasons. First, previous studies rely on information on cross-country Mincerian returns to education and define their skill-premium to be the return to completing certain level of schooling, which could either be elementary school, secondary school, or university. As discussed by the recent empirical labor literature, this approach has the disadvantage that in the last three decades, a period often characterized by an acceleration in technological progress, educational attainment has lost explanatory power in wage regressions. At the same time, the explanatory power of occupations in accounting for wage differences across workers has significantly increased in this period of time, not only in developed, but also in developing countries [10].

Second, there are other characteristics that affect worker's skills besides their formal education, like their own innate ability to perform different tasks, their general health status, and the learning that might be acquired on the job, among others. This set of characteristics are better captured by observed labor income rather than by expected average returns to

[10] For example, Acemoglu and Autor (2011a) document that even though the university/ secondary school wage premium has monotonically increased since the 1970s in the US, these changes in wage levels and the distribution of wages have been accompanied by systematic, non-monotone shifts in the composition of employment across occupations, with rapid simultaneous growth of both high education high wage occupations and low education, low wage occupations in the United States and the European Union.

education. Third, the development literature has not reached an agreement on what level of educational attainment should be considered to be the minimum in order to classify workers as skilled. The main implication of choosing different educational attainment levels is that the higher the educational attainment level required to classify workers as skilled, the larger the variability of the resulting aggregate human capital stock across countries, increasing its explanatory power in development accounting exercises.

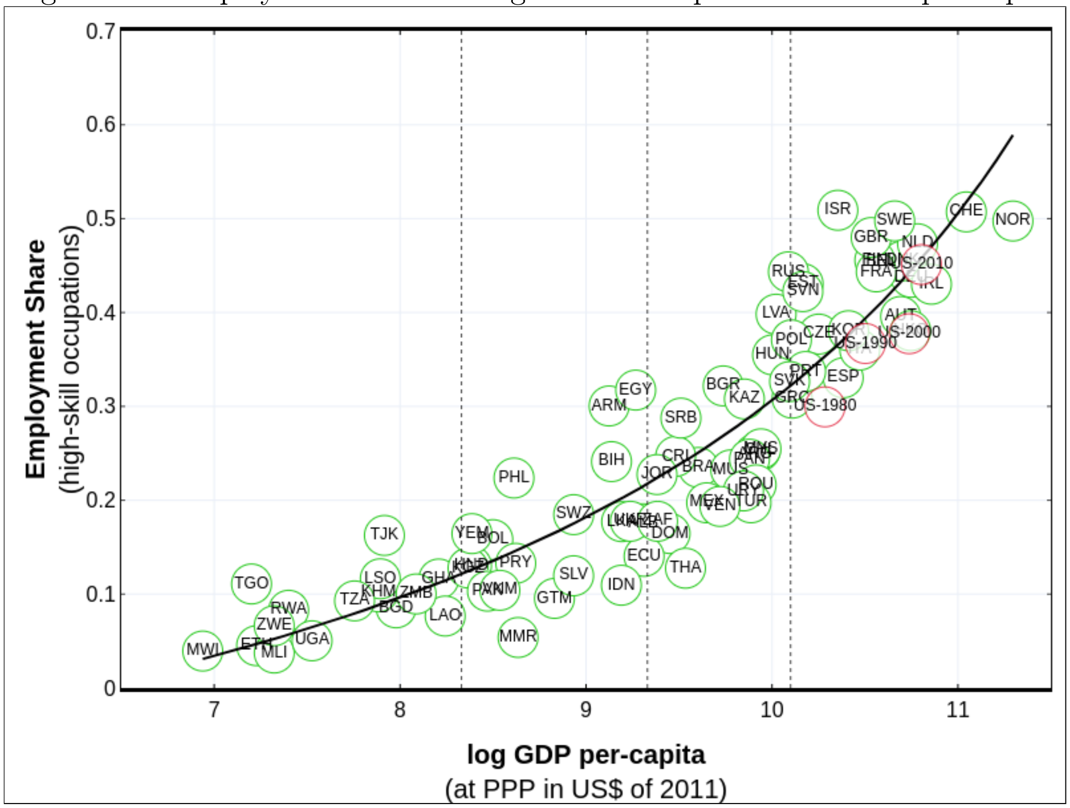
1.3 Empirical Results.

I begin by studying the evolution of occupational employment at different development levels. To that end, I calculate the fraction of total employment in high-skill occupations for the countries in my sample and plot them against their corresponding GDP per capita. In addition, I fit a regression for the employment share in high-skill occupations on a fractional polynomial model in GDP per capita, which is represented by the solid black line in the figure [11]. The vertical dashed lines separate the sample into development quartiles. The results are presented in Figure 1.1 below.

The main message from Figure 1.1 is that employment in high-skill occupations rises as countries grow richer. Quantitatively, the average share of employment in high-skill occupations rises from 8.5 percent to 18.1, 27.5, and 42.8 respectively as we move from the first to the fourth development quartile. When compared to educational attainment data, the share of employment in high-skill occupations is higher than the average fraction of workers with university complete (1.98 percent) and lower than the proportion of individuals with secondary education complete (15.8) at the first development quartile. However, it grows at

[11] The best fitting model is a one-dimensional polynomial of the form $\left(\frac{L_{hs,c}}{L_c}\right) = \beta_0 + \left(\frac{\beta_1}{log(y_c) - y_0}\right) + \epsilon_c$. Estimated coefficients are both jointly and individually statistically significant at the one percent level and given by: $\beta_0 = -0.36$, $\beta_1 = -2.90$, and $y_0 = 14.35$.

Figure 1.1: Employment Share in High-Skill Occupations vs GDP per capita.

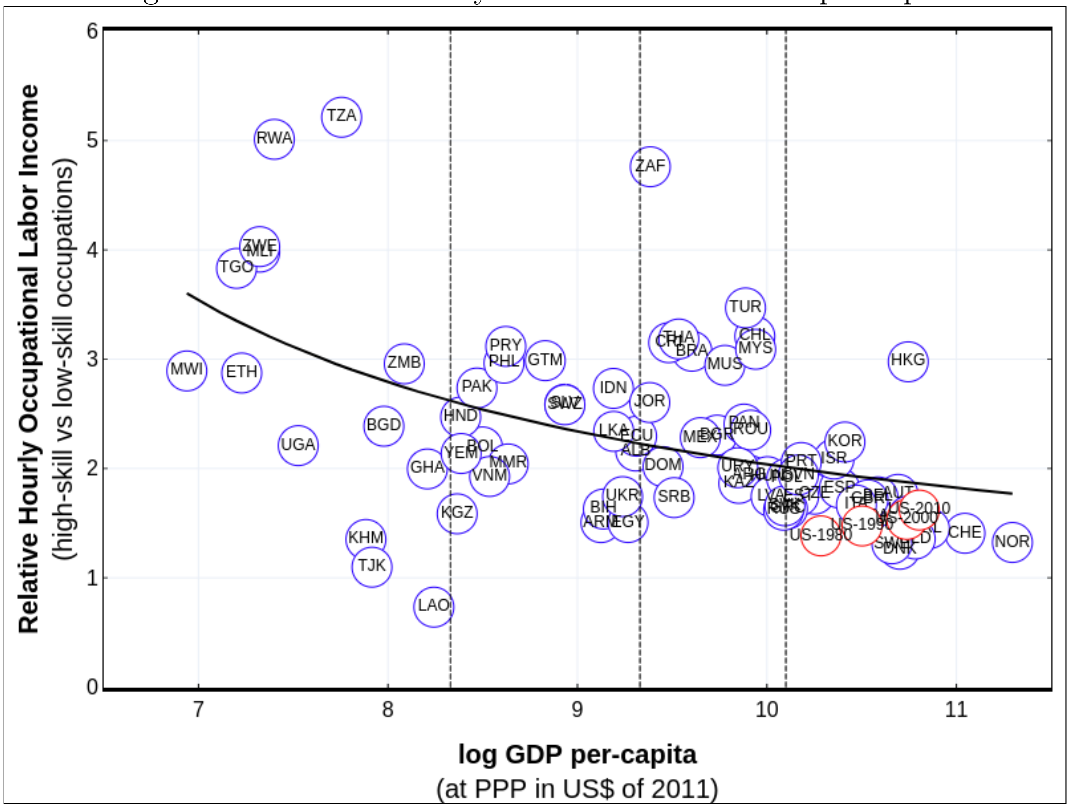


a faster speed that both measures with development, reaching an average of 43.5 percent for countries in the highest quartile of development, while the average fractions of individuals with secondary and university complete total 38.8 and 13.7 percent, respectively.

I proceed to study the behavior of the occupational skill-premium across the development spectrum. Therefore, I plot each country's occupational skill-premium and GDP per capita in Figure 1.2 below. As in the case of the employment shares, the solid black line in the figure represents the best fitting fractional polynomial model [12] and the vertical dashed lines separate the sample into development quartiles.

[12] The best fitting model is a one-dimensional polynomial of the form $\left(\frac{w_{hs,c}}{w_{ls,c}}\right) = \beta_0 + \left(\frac{\beta_1}{log(y_c) - y_0}\right) + \epsilon_c$. Estimated coefficients are both jointly and individually statistically significant at the one percent level and given by: $\beta_0 = 0.47$, $\beta_1 = 9.61$, and $y_0 = 3.86$.

Figure 1.2: Relative Hourly Labor Income vs GDP per capita.

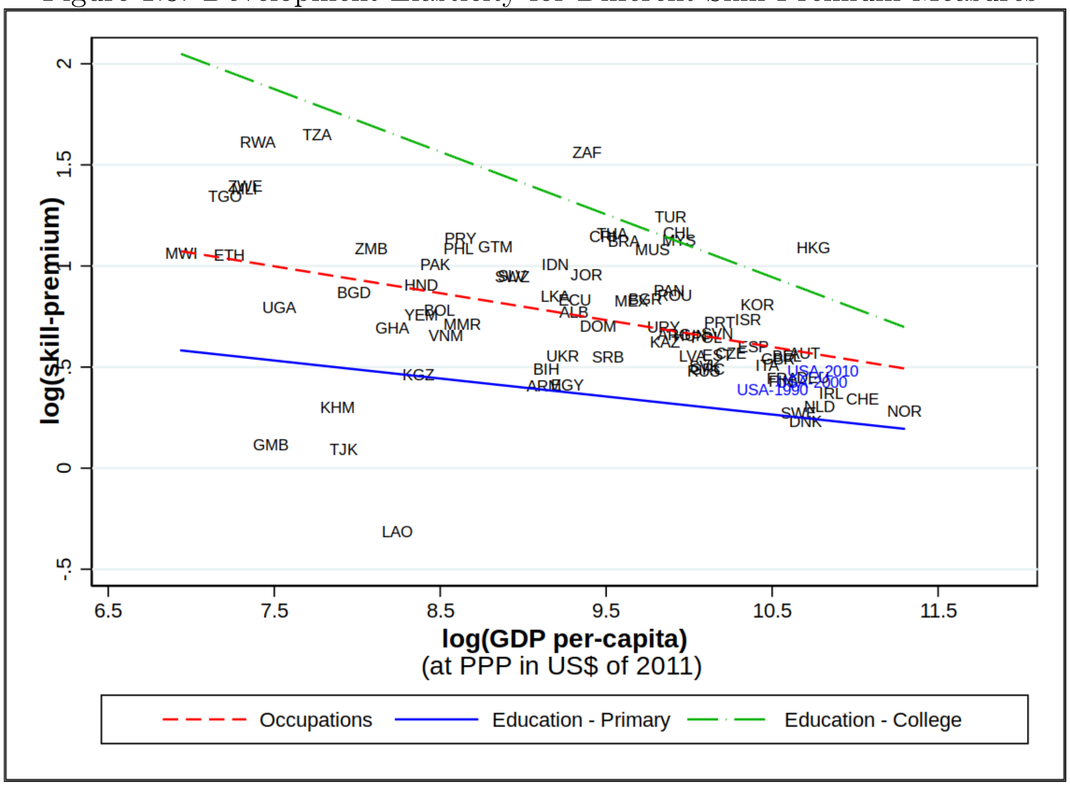


The main message from Figure 1.2 is that there is a negative relationship between the occupational skill-premium and GDP per capita. As the regression line shows, this relationship is fairly non-linear, exhibiting a steeper decline at lower development levels. On average, the occupational skill-premium falls from 2.5, to 2.3, 2.1 and 1.6 as we move from the first to the fourth quartile of my sample. Compared to the skill-premium measures based on Mincerian returns to education used by Caselli and Coleman, these numbers lie in between the estimates that takes as high-skill workers those with primary complete and the one that defines as high-skill workers those with university complete, respectively. The occupational skill-premium lies below the educational skill-premium metric that takes workers with university or above as the high-skill group [13]. Figure 1.3 below presents a comparison between

[13] When workers with primary education or higher are defined as high-skill Caselli and Coleman's measure falls from 1.7 to 1.5, 1.4 and 1.3 as we move from the first to the fourth quarter of income per capita in their sample. When secondary complete is taken as the minimum level of school attainment to be considered high-

the occupational skill-premium and the measures of skill-premia based on educational attainment, assuming that skill premium elasticities are constant across development for all measures. I expand on the non-linearities of the skill-premium elasticity in the following section.

Figure 1.3: Development Elasticity for Different Skill-Premium Measures



1.3.1 Occupational Skill-Premium Non-Linearities Across Development.

I here study in further detail the non-linear relationship between the occupational skill-premium and GDP per capita. In order to quantitatively asses these non-linearities in a more tractable framework than the one suggested by the fractional polynomial fit, I estimate

skill, there measure falls from 3.7 to 2.7, 2.4 and 2.0, respectively. If the educational attainment threshold is university complete instead, the corresponding numbers are 9.8, 4.7, 4.0, and 2.9

the skill-premium development elasticity by running the following linear regression

$$log\left(\frac{w_{hs,c}}{w_{ls,c}}\right) = \beta_0 + \beta_1 \cdot log\left(y_c\right), \tag{1.1}$$

and compare it with a model where I let the elasticity vary across development quartiles, by fitting the following OLS regression to the data

$$log\left(\frac{w_{hs,c}}{w_{ls,c}}\right) = \beta_0 + \beta_1 \cdot log\left(y_c\right) + \sum_{q=2}^{4} 1_{[c \in q]} \cdot \beta_q \cdot log\left(y_c\right). \tag{1.2}$$

In Equations 1.1 and 1.2 , $log\left(\frac{w_{hs,c}}{w_{ls,c}}\right)$ is the natural logarithm of the occupational skill-premium in country c and $log\left(y_c\right)$ is the natural logarithm of real GDP per capita at PPP in constant US$ of 2011 for country c. In Equation 1.2, q is quartile indicator, the higher the more developed a country is, and the corresponding quartile indicator functions take the value of one $1_{[c \in q]}$ if country c belongs to development quartile q.

Equation 1.1 is standard and does not require further discussion. Equation 1.2, β_1 measures the occupational skill-premium elasticity with respect to GDP per capita for countries in the first development quartile, which I choose to be the base. In this model β_2, β_3, and β_4 measure the change in the base elasticity as countries move to higher development levels, represented by the second, third, and fourth quartiles, respectively. For these coefficients, classical statistical tests of significance (i.e. t-tests) show if the estimated elasticity for the corresponding quartile is statistically different from the base one.

The estimation results for Equations 1.1 and 1.2 are presented under the names of Model (1) and Model (2) in Table 1.2 below.

Table 1.2: Development Elasticity of the Occupational Skill-Premium
(quantitative assessment of non-linearities)

Variables	**Model (1)** $log\left(\frac{w_{hs}}{w_{ls}}\right)$	**Model (2)** $log\left(\frac{w_{hs}}{w_{ls}}\right)$	**Model (3)** $log\left(\frac{w_{hs}}{w_{ls}}\right)$
$log\,(y_c)$	-0.127***	-0.391***	-0.416***
	(0.035)	(0.118)	(0.097)
$1_{[c\in 2]} \cdot log\,(y_c)$	-	0.051**	0.054***
	-	(0.019)	(0.018)
$1_{[c\in 3]} \cdot log\,(y_c)$	-	0.076***	-
	-	(0.006)	-
$1_{[c\in 4]} \cdot log\,(y_c)$	-	0.072*	-
	-	(0.033)	-
$1_{[c\in 3 \vert c\in 4]} \cdot log\,(y_c)$	-	-	0.080***
	-	-	(0.025)
Constant	1.949***	3.906***	4.101***
	(0.329)	(0.923)	(0.754)
Observations	80	80	80
R-squared	0.146	0.254	0.253
Adjusted R-squared	0.135	0.214	0.223
Prob >F	0.0005	0.0002	0.0001

* Standard errors in parentheses (*** p<0.01, ** p<0.05, * p<0.1).

Model (1) confirms that there is a statistically significant negative relationship between GDP per capita and the skill-premium. The estimated elasticity is -0.12 and is statistically significant at the one percent level. The model implies that the skill-premium falls from 2.4 to 1.5 when countries move from the median GDP per capita in the first quartile to the corresponding one in the fourth quartile.

Model (2) shows the skill-premium elasticity with respect to GDP per capita is not only negative, but also that it declines in absolute terms as countries move from the first to the fourth development quartile. Measured elasticities are -0.339, -0.294, -0.275, and -0.279 for the first, second, third, and fourth quartile, respectively, and are all statistically significant, at least at the ten percent level. The model implies that the skill-premium falls from 2.20 to 1.65 when countries move from the average GDP per capita in the first quartile to the average income per capita in the fourth quartile.

A statistical test of joint linear restrictions rejects the nulls that $\beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$, and $\beta_1 = \beta_2 = \beta_3 = \beta_4$, both at the 1 percent level, and that $\beta_2 = \beta_3 = \beta_4 = 0$ at the 5 percent level. However, the null that $\beta_2 = \beta_3 = \beta_4$ can not be rejected at ten percent level, which favors the alternative that at least one pair of these coefficients are equal.

As a consequence, I estimate four additional regressions that capture all possible combinations of equality between these three coefficients. The best specification, as measured by the highest adjusted R^2 attained, is given by the model where $\beta_1 \neq \beta_2 \neq (\beta_3 = \beta_4)$, which I present in third column of Table 1.2 and name Model (3).

Under the specification given by Model (3), the skill-premium elasticity falls from -0.362 in the bottom quartile of development to -0.314 in the second quartile, stabilizing at -0.293 in the third and fourth quartiles. These elasticities are all statistically significant at the one percent level and the model predicts a decline in the skill-premium from 2.21 to 1.65 as we move from the median GDP per capita in the poorest to quartile to the median GDP per capita in the richest quartile.

As a summary, I find that there is a negative, statistically significant relationship between the occupational skill-premium and development. This relationship is fairly non-linear, exhibiting a steeper decline at the first development quartile. The best model specification to measure these non-linearities implies an predicted cccupational skill-premium that falls from 2.21 to 1.65 as we move from the poorest to the richest development group [14].

[14]I estimate the same regressions using the data on GDP per-worker and educational skill-premium provided by Caselli and Coleman and find a statistically significant negative elasticity for their three measures. In this case, the best specifications throw skill-premium elasticities that do not vary with development. Es-

1.3.2 Robustness.

In this section I perform sensitivity exercises to asses the robustness of the quantitative results presented in Section 1.3.1 above. In the first set of robustness checks I explore to what extent my results are driven by the criterion used to group occupations into broad categories and the role of extreme occupational skill-premium values. The second group of exercises analyze the quantitative relevance of institutions, how the results change if I exclude female workers from my sample, and the role of hours worked across development.

The first set of checks is to guarantee that my results are not driven by the criterion used to define high-skill occupations or by countries with relatively high or low occupational skill-premium values. The spirit of the second group of exercises is to disentangle the underlying mechanisms that drive the decline in the occupational skill-premium across development so that at the time of writing a model one can focus on frameworks that shed light on those forces.

Appendix A.4 presents a detailed description on how this robustness checks are performed. As a summary, neither changing the criterion to classify occupations into high- and low-skill, nor excluding extreme values or adding institutional controls produce any major qualitative changes to the results, with the best model that fits the data still being the one in which the occupational skill-premium falls with development until it reaches the third GDP per capita quartile. Quantitatively, the estimated elasticities are in the same range as those reported in Table 1.2 above, increasing modestly in absolute terms if institutional controls are added or if the lowest median wage occupation in the high-skill group is included in the low-skill group. Shifting occupations in the margin of the low-skill group to the high-skill

timated elasticities are -0.32, -0.20, and -0.08, depending is the schooling threshold for high-skill workers is university, secondary, or primary complete, respectively.

group reduces the predicted skill-premium and the corresponding elasticities, but again, only modestly.

Moving to the second set of exercises, the exclusion of women does not produce any major qualitative or quantitative changes in the results. The best model is still given by the one were the occupational skill-premium elasticity declines with development until countries reach the third quartile of GDP per capita. The estimated elasticities are fairly similar as those presented in Table 1.2. This robustness check suggests that, for example, one can safely discard frameworks that exploit the role of women in the labor force and their attachment to the labor market when the object of study is the decline in occupational skill-premium across development.

Computing the occupational skill-premium without controlling for hours worked leads to similar quantitative and qualitative results. The best statistical fit is still given by the model where elasticities decline with development. When compared to the estimated that control for hours worked, measured elasticities decline marginally in absolute terms, but are still in line with those presented in Table 1.2. This is mainly due to the fact that hours worked exhibit a higher decline in low- than in high-skill occupations across development. As in the previous case, this suggests that the main driver of the decline of the occupational skill-premium is not hours worked and one can abstract from models whose main focus is on the intensive margin of labor.

1.4 Model

In what follows I build a stylized model that is useful to understand the labor market structure and evolution through development and allows me to understand the main mechanisms that account for the observations presented in Section 1.3 above. The ultimate goal is to use the model to understand the determinants of relative labor productivity in high- and low-skill occupations as countries grow richer.

1.4.1 Environment

The economy is populated by a continuum of individuals of measure one. They are endowed with a unit of time and a pair of occupational-specific labor productivities $z = \{z_l, z_h\}$, where z_h (z_l) represents a realization of a worker's labor productivity in high-skill (low-skill) occupations. Labor productivity is jointly distributed with cumulative distribution function $G(z_h, z_l)$, and $g(z_h, z_l)$ represents the corresponding probability density function. Considering that in my empirical analysis I find that hours worked are not the main driving force of the decline in the occupational skill-premium we observe in the data, I assume that workers provide their time to the labor market fully and inelastically.

Output is produced by combining occupational labor services, according to the technology

$$y = \left[(A_l L_l)^{\left(\frac{\sigma-1}{\sigma}\right)} + (A_h L_h)^{\left(\frac{\sigma-1}{\sigma}\right)}\right]^{\left(\frac{\sigma}{\sigma-1}\right)} \tag{1.3}$$

where $L_l = \int_{\mathcal{L}} z_l \left(\int_0^{+\infty} g(z_h, z_l) dz_h\right) dz_l$ and $L_h = \int_{\mathcal{H}} z_h \left(\int_0^{+\infty} g(z_h, z_l) dz_l\right) dz_h$ are total labor output in low- and high-skill occupations, σ represents the elasticity of substitution between skill types, A_l and A_h are occupation-specific productivity parameters that transform labor outputs into labor services, and $\mathcal{L}$ and $\mathcal{H}$ denote the set of individuals who work in low- and high-skill occupations, respectively.

Assuming that the aggregate technology is operated by a cost minimizing firm that acts competitively in both labor markets, relative wages are given by

$$\left(\frac{w_h}{w_l}\right) = \left(\frac{L_h}{L_l}\right)^{-\left(\frac{1}{\sigma}\right)} \left(\frac{A_h}{A_l}\right)^{\left(\frac{\sigma-1}{\sigma}\right)}. \tag{1.4}$$

Fixing the relative supply of skilled labor, if labor types are are substitutes in production ($\sigma > 1$), economic development processes characterized by skilled-biased technological change ($\uparrow A_h \ > \uparrow A_l$) lead to an increase in relative efficiency wages. On the other hand, if labor services are complements in production ($\sigma < 1$), relative efficiency wages grow with development if technological progress is unskilled labor biased ($\uparrow A_l \ > \uparrow A_h$). In the Cobb-Douglas case ($\sigma = 1$), irrespective of its nature, technological progress has a neutral effect on relative wages.

1.4.2 A Basic Roy Model for The Labor Market.

Assume markets are competitive and workers can freely select their occupation. Let w_h and w_l represent wages per efficiency unit of labor in high- and low-skill occupations. As is standard in Roy models, individuals choose to work in high-skill occupations if their labor income W_h^i is higher than in low-skill occupations W_l^i, or

$$W_h^i > W_l^i$$

$$w_h z_h^i > w_l z_l^i.$$

Re-arranging terms

$$\frac{z_h^i}{z_l^i} > \left(\frac{w_l}{w_h}\right),$$

and agents choose to work in high-skill occupations if their comparative advantage in them is higher than the inverse of relative wages in high- and low-skill occupations. For very low high-to-low skill occupations relative wage levels, only workers with a very high comparative advantage in high-skill occupations choose them. As wages in high-skill occupations grow with respect to those in low-skill occupations workers with lower comparative advantage in high-skill occupations choose to work in them.

In this context, the set of workers who choose high-skill occupations is

$$\mathcal{H} = \Big\{ i \in [0,1] : w_h z_h^i > w_l z_l^i \Big\},$$

and their corresponding effective labor supply is

$$\begin{aligned} L_h &= \int_{\mathcal{H}} z_h \left(\int_0^{+\infty} g(z_h, z_l) dz_h \right) dz_l \\ &= \int_0^{+\infty} z_h \int_0^{+\infty} g(z_h, z_l \wedge w_h z_h > w_l z_l) dz_l dz_h \\ &= \pi_h \int_0^{+\infty} z_h \int_0^{+\infty} g(z_h, z_l \mid w_h z_h > w_l z_l) dz_l dz_h \\ &= \pi_h \int_0^{+\infty} z_h \int_0^{z_h \left(\frac{w_h}{w_l}\right)} g(z_h, z_l) dz_l dz_h = \pi_h \cdot \mathbb{E}\big(Z_h | w_h z_h > w_l z_l\big). \end{aligned} \tag{1.5}$$

Similarly, the effective labor supply of workers in low-skill occupations is

$$L_l = \pi_l \cdot \mathbb{E}\big(Z_l | w_h z_h \leq w_l z_l\big). \tag{1.6}$$

where π_h denotes the fraction of workers in high-skill occupations, given by

$$\pi_h = Prob\big(w_h Z_h > w_l Z_l\big) = \int_{-\infty}^{+\infty} \int_0^{z_h \left(\frac{w_h}{w_l}\right)} g(z_h, z_l) dz_l dz_h, \tag{1.7}$$

and $\pi_l = (1 - \pi_h)$.

To ease notation, in what follows I call $\bar{z_h} = \mathbb{E}\big(Z_h | w_h z_h > w_l z_l\big)$ and $\bar{z_l} = \mathbb{E}\big(Z_l | w_h z_h \leq w_l z_l\big)$. Thus, from now on $L_h = \pi_h \bar{z_h}$ and $L_l = \pi_l \bar{z_l}$.

As we can see from Equation above, a rise in the high-to-low skill wage ratio $\left(\frac{w_h}{w_l}\right)$ leads to a rise in the faction of workers who choose high-skill occupations. This is as a consequence of the standard selection mechanism described above, and is reflected in an increase in the upper limit of integration in the inner integral of Equation . More importantly, this occurs independently of the properties of the joint skill distribution $g(z_h, z_l)$.

Continuing with the characterization of the main model objects of interest, let $M_h\left(z_h\right)$ denote the cumulative labor productivity distribution function conditional on workers selecting high-skill occupations and $m_h\left(z_h\right)$ its corresponding pdf. The former is given by

$$\begin{aligned} M_h\left(z\right) &= Prob\big(Z_h \leq z | w_h z_h > w_l z_l\big) \\ &= \frac{Prob\big(Z_h \leq z, w_h z_h > w_l z_l\big)}{Prob\big(w_h z_h > w_l z_l\big)} \\ &= \left(\frac{1}{\pi_h}\right) \int_0^z \int_0^{z_h \left(\frac{w_h}{w_l}\right)} g(z_h, z_l) dz_l dz_h \\ &= \left(\frac{1}{\pi_h}\right) \int_0^z g\left(z_h, z_h \left(\frac{w_h}{w_l}\right)\right) dz_h. \end{aligned}$$

Similarly, the corresponding labor productivity cdf conditional on choosing low-skill occupations is

$$M_l\left(z\right) = \left(\frac{1}{\pi_l}\right) \int_0^z g\left(z_l \left(\frac{w_l}{w_h}\right), z_l\right) dz_l.$$

These two distributions are the relevant empirical objects of interest, in consideration of the fact that the researcher observes certain characteristics of workers conditional on their occupational decision. Finally, average labor income in high- and low-skill occupations are represented by

$$\bar{W_h} = w_h \cdot \mathbb{E}\big(Z_h | w_h z_h > w_l z_l\big) = w_h \cdot \bar{z_h},$$

and

$$\bar{W_l} = w_l \cdot \mathbb{E}\big(Z_l | w_h z_h \leq w_l z_l\big) = w_l \cdot \bar{z_l}.$$

As a consequence, the model's counterpart for the occupational skill-premium, or the ratio of average labor income in high- and low-skill occupations presented in Figure 1.2 is

$$\left(\frac{\bar{W_h}}{\bar{W_l}}\right) = \left(\frac{w_h}{w_l}\right) \frac{\mathbb{E}\big(Z_h | w_h z_h > w_l z_l\big)}{\mathbb{E}\big(Z_l | w_l z_l \geq w_h z_h\big)} = \left(\frac{w_h}{w_l}\right) \left(\frac{\bar{z_h}}{\bar{z_l}}\right) \tag{1.8}$$

Through the lens of this framework, and assuming that $\sigma > 1$, economic development processes characterized by skilled-biased technological change ($\uparrow \frac{A_h}{A_l}$) lead to a rise in the efficiency wage of high-skill occupations relative to its low-skill occupations counterpart $\left(\frac{w_h}{w_l}\right)$, which lowers the comparative advantage that workers require to choose high-skill occupations $\left(\frac{z_h}{z_l}\right)$ and drives the fraction of workers in high-skill occupations π_h up.

At this point, one can not give a conclusive answer about the behavior of the relative labor income ratio $\left(\frac{\bar{W_h}}{\bar{W_l}}\right)$, since it might be possible that the rise in efficiency wages $\left(\frac{w_h}{w_l}\right)$ lead to either an increase or a decline in relative mean labor productivities $\left(\frac{\bar{z_h}}{\bar{z_l}}\right)$, which can potentially augment, offset, or more than offset the initial effect of relative wages.

Simply stated, in order to know how average labor productivity in high- $(\bar{z_h})$ and low-skill $(\bar{z_l})$ occupations react to an increase in the share of individuals choosing high-skill occupations, one needs to know the marginal worker's absolute productivity levels in high- and low-skill occupations and how they are related to the population means. This can not be fully described without further knowledge on the properties of the joint skill distribution in the population, denoted in the model by $g(z_h, z_l)$.

1.4.3 Absolute and Relative Occupational Labor Productivity and The Properties of The Joint Skill Distribution.

Having described the forces in the model that lead to an increase in the fraction of workers in high-skill occupations as a response to skilled-biased technical change, or higher relative skilled labor efficiency, I proceed to study in further detail the model mechanisms that allow me to interpret the decline in the occupational skill-premium I observe in the data.

To illustrate the role that the joint distribution of skills plays in occupational choice and in determining the levels of occupational labor productivity, consider the case were employment in high-skill occupations is low. In such scenario, the relative wage in high-skill occupations is low enough that only workers with a relatively high comparative advantage in these jobs choose them. This might be either because they have high absolute advantage in High-Skill occupations and low absolute advantage in low-skill occupations, or because they have low absolute advantage in high- and low-skill occupations but they are amongst the individuals with lowest absolute advantage in low-skill occupations in the population, or because they are endowed with high absolute advantage in high- and low-skill occupations but they are amongst the individuals with highest absolute advantage in high-skill occupations in the population [15].

[15] Heckman and Honoré (1990) explore these three alternatives and characterize the properties of the Roy model and its identification for the case in which the distribution of skill differences are log-concave.

In the first case there is positive selection due to unobservable ability both in high- and in low-skill occupations. In this situation, the joint skill distribution is such that, for workers in both occupational groups, their absolute skill level is higher than the corresponding population average. As a consequence, as the fraction of workers in high-skill occupations rises and the required comparative advantage to select them declines, workers of lower absolute advantage in high-skill jobs now enter them, which drives down average labor productivity in these occupations. For low-skill occupations instead, it is the workers with lower absolute advantage who leave this group, increasing average labor productivity.

The second case features positive selection in high-skill occupations and negative selection in low-skill occupations. In this situation, the individuals who move from low-skill to high-skill occupations have an absolute advantage that is smaller than the average for those working in high-skill jobs and are the highest ability individuals in low-skill jobs before switching occupations. Thus, average labor productivity falls both in High- and low-skill occupations.

The third case is when there is negative selection in high-skill occupations and positive selection in low-skill occupations. In this context, an increase in the fraction of workers in high-skill occupations leads to a rise in average labor productivity both in high- and low-skill occupations. The former is straightforward, as the few workers in high-skill occupations before the increase in wages were those with lowest absolute advantage in the population. The fact that the average skill of workers in low-skill occupations rises is as a result of the lowest ability individuals in low-skill jobs switching occupations [16].

[16] Heckman and Honoré rule out the possibility for negative selection in both occupational groups since it requires the covariance of the joint skill distribution to be larger than the variances for both marginals, which leads to a variance-covariance matrix that is not positive semi-definite.

It should be clear at this point that identifying the joint distribution of skills in the population is fundamental in order to understand what are the driving forces behind the decline in the ratio of average occupational labor income as countries develop.

In order to identify the underlying properties of the joint distribution of skills in the population, I proceed as follows. For tractability purposes, I restrict my attention to parametric distributions. This simplifies the identification problem to finding a family of distributions that fit the data reasonably well in the first place, to in turn proceed with the corresponding parameter estimation.

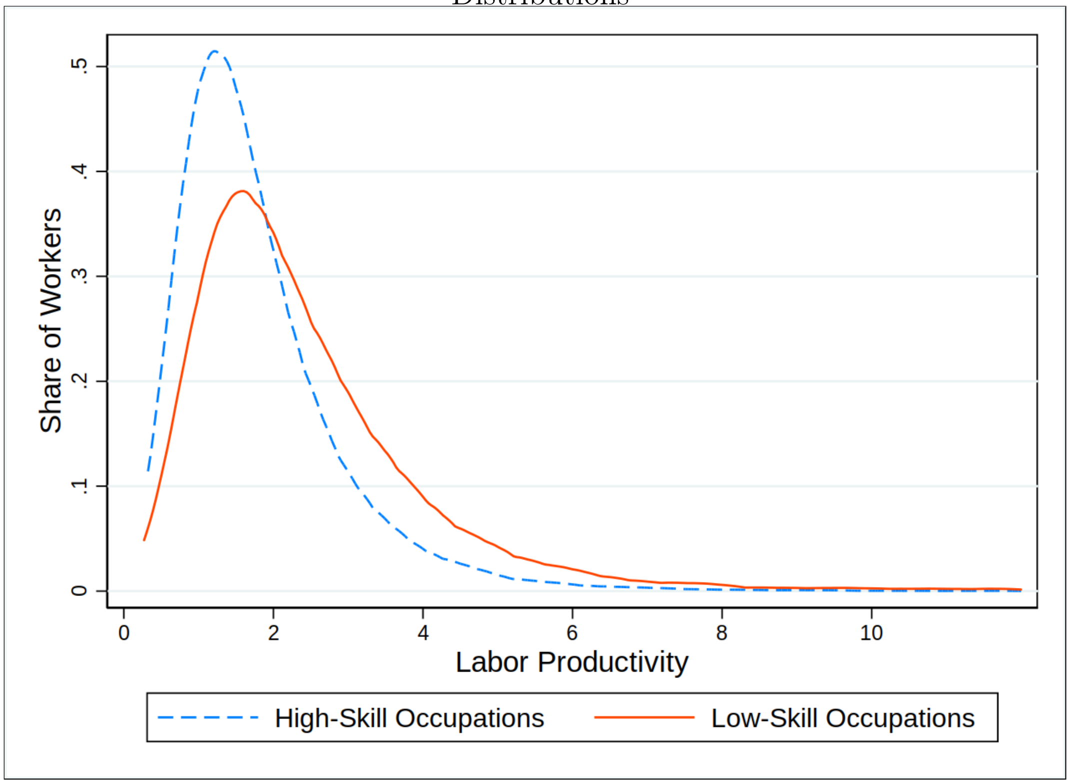
I discipline my distributional choice by looking at the empirical labor productivity distributions conditional on workers occupational choice ($M(z_h)$ and $M(z_l)$). To do so, I use micro-data on labor income, hours worked, workers' main occupation and a set of demographic characteristics from IPUMS International for the United States in 2010 to compute hourly labor income by occupation at the individual level [17]. Separating the component of workers' labor income that corresponds to their productivity or human capital from the wage per efficiency unit of labor requires an identifying assumption [18]. I choose to normalize the labor productivity of white male workers with no experience and no formal education in High and low-skill occupations to unity. As a result, the average labor income for the individuals in these groups represent the wage per efficiency unit of labor in High and low-skill Occupations, respectively. Dividing the labor income of workers in High and low-skill occupations by their corresponding efficiency wages gives me a measure of the productivity of workers in these two occupations. The Kernel density estimates of the labor productivity

[17] IPUMS International microdata for the US is based on Census data. I use IPUMS International over IPUMS USA because in the former workers occupations are presented using the ISCO classification, as in the ILO databse. Data handling and issues are explained in further detail in Appendix A.6.

[18] This is standard in the literature. See, for example, Buera et al. (2018).

distributions by broad occupational group are presented in Figure 1.4 below.

Figure 1.4: Kernel Density Estimation of the Occupational Labor Productivity Distributions



The shape of the empirical densities suggest that the Log-Normal and Frechet cases are good candidates [19]. To decide between these two parametric families, I impose the additional restriction that, for the distribution chosen, the model should be able to reproduce a decline in relative average labor income $\left(\frac{\bar{W}_h}{\bar{W}_l}\right)$, like I document in Section 1.2 above. Two comments are in order. First, notice that this restriction does not entail quantitatively targeting any data moment. It only asks the model to qualitatively reproduce the behavior of relative occupational labor income across development. Second, by imposing this restric-

[19]A Kolmogorov-Smirnov test does not reject the hypotheses that the occupational labor productivity distributions are both distributed Log-Normally or both distributed Frechet at the one percent level. On the other hand, other long tailed distributions like the Beta and Pareto cases are rejected at the ten percent level.

tion I am not ruling out any of the types of selection discussed above. For example, under bivariate Log-Normally distributed skills one can potentially obtain the decline in relative labor income in the three cases discussed above. The selection type that the model features ultimately depend on the estimated values of the variances of the marginal distributions and the covariance between them [20]. However, as I show in Appendix A.5, in the Log-Normal case one can not obtain a labor income ratio that monotonically declines with development.

As it will become more clear in Section 1.4.5 below, a monotonically declining ratio of labor income between high- and low-skill occupations can be obtained in the Frechet case under correlated marginal distributions. Before exploring the role of dependence, I characterize the properties of the main model objects under independently distributed Frechet skills below, which are qualitatively preserved once dependence is considered.

1.4.4 The Roy Model Under Independent Frechet Skill Distributions.

I here analyze some basic properties of the model under the assumption that occupational abilities are drawn from independent Frechet distributions. In order to obtain analytical results I present the common shape parameter case. The properties presented below are preserved under shape parameters that differ across marginals.

Proposition 1. *Assume skills are drawn from two independent* $Frechet(S_j, \theta) = e^{-S_j z_j^{-\theta}}$ *distributions, where* S_j *is an occupation-specific scale parameter governing the overall level of the productivity draws, and* θ *is a shape parameter that controls the degree of variation in the distribution.*

[20] Case 1: $\sigma_h > \sigma_{hl}$ and $\sigma_l > \sigma_{hl}$; more likely under $\sigma_{hl} < 0$. Case 2: $\sigma_h > \sigma_{hl}$ and $\sigma_l < \sigma hl$; $\sigma hl > 0$. Case 3: $\sigma h < \sigma hl$ and $\sigma l > \sigma hl$; $\sigma hl > 0$

1. *The share of workers in high-skill occupations is given by*

$$\pi_h = Prob\big(w_h Z_h > w_l Z_l\big) = \frac{S_h}{\left(S_h + S_l \left(\frac{w_h}{w_l}\right)^{-\theta}\right)} \ ,$$

2. *The skill distribution conditional on workers selecting high-skill occupations is given by*

$$M_h(z) = e^{-\left(S_h + S_l\left(\frac{w_h}{w_l}\right)^{-\theta}\right)z^{-\theta}}.$$

3. *The skill distribution conditional on workers selecting low-skill occupations is*

$$M_l(z) = e^{-\left(S_l + S_h\left(\frac{w_h}{w_l}\right)^{\theta}\right)z^{-\theta}}.$$

4. *Average labor productivity in high- and low-skill occupations are given by*

$$\bar{z_h} = \mathbb{E}\big(Z_h | w_h z_h > w_l z_l\big) = \left(S_h + S_l \left(\frac{w_h}{w_l}\right)^{-\theta}\right)^{\frac{1}{\theta}} \Gamma\left(1 - \frac{1}{\theta}\right),$$

and

$$\bar{z_l} = \mathbb{E}\big(Z_l | w_l z_l \geq w_h z_h\big) = \left(S_l + S_h \left(\frac{w_h}{w_l}\right)^{\theta}\right)^{\frac{1}{\theta}} \Gamma\left(1 - \frac{1}{\theta}\right),$$

where $\Gamma\left(1 - \frac{1}{\theta}\right)$ *is the Gamma function evaluated at* $\left(1 - \frac{1}{\theta}\right)$*.*

5. *The ratio of occupational average labor incomes is*

$$\frac{\bar{W_h}}{\bar{W_l}} = \left(\frac{w_h}{w_l}\right) \frac{\mathbb{E}\big(Z_h | w_h z_h > w_l z_l\big)}{\mathbb{E}\big(Z_l | w_l z_l \geq w_h z_h\big)} = \left(\frac{w_h}{w_l}\right) \frac{\left(S_h + S_l\left(\frac{w_h}{w_l}\right)^{-\theta}\right)^{\frac{1}{\theta}}}{\left(S_h\left(\frac{w_h}{w_l}\right)^{\theta} + S_l\right)^{\frac{1}{\theta}}} = 1.$$

Proof. See Appendix A.5. $\square$

Item 1 fully characterizes the fraction of workers in high-skill occupations. An increase in the scale parameter of the high-skill innate ability distribution (S_h) rises the share of workers in high-skill occupations. Intuitively, fixing S_l and $(\frac{w_h}{w_l})$, the higher S_h is, the higher worker's comparative advantage in high-skill occupations is, and the more likely they are to select them. The opposite holds for the scale parameter of the low-skill innate ability distribution, S_l. A rise in the skill-premium $\left(\frac{w_h}{w_l}\right)$ leads to an increase in the fraction of workers in high-skill occupations, since now it is more likely that workers with a lower comparative advantage in high-skill occupations will choose to work in them.

Items 2 and 3 show how the selection of workers into occupations endogenously affects the scale parameter of the conditional skill distributions, given by $\left(S_h + S_l\left(\frac{w_h}{w_l}\right)^{-\theta}\right)$ and $\left(S_l + S_h\left(\frac{w_h}{w_l}\right)^{\theta}\right)$, in the high- and low-skill case, respectively. Focus on the high-skill case for a moment. An increase in the marginal effective wage ratio $(\frac{w_h}{w_l})$ lowers the scale parameter of the high-skill ability distribution. Thus, after an increase in relative wages we are more likely to observe workers with lower ability levels in high-skill occupations. This is due to the fact that the smaller required comparative advantage to work in high-skill occupations after a rise in relative wages is associated with a smaller absolute advantage in these type of jobs, which lowers their average ability level. This effect moves the conditional skill distribution for workers in high-skill occupations up and to the left. The opposite holds for low-skill occupations.

The intuition above is the key to understand the results in item 4, which show that average labor productivity in high-skill (low-skill) occupations is decreasing (increasing) in the efficiency wage ratio.

Item 5 in Proposition 1.4.5 shows that in the case of independent Frechet distributions with common shape parameter the ratio of average occupational labor income is constant and independent of the wage ratio ($\frac{w_h}{w_l}$). This is due to the fact that the increase in average labor income in high-skill occupations after a rise in the skill-premium is exactly compensated by the decline in average labor productivity in high-skill occupations and the increase in average labor productivity in low-skill occupations [21]. In the following section I show how an increase in the wage premium can lead to a decline in the ratio of average occupational labor incomes if the independence assumption is relaxed.

1.4.5 The Roy Model Under Joint Frechet Skill Distributions.

In this section I show how the observed increase in the fraction of workers in high-skill occupations and decline in the high-to-low-skill average labor income ratios as countries develop can be rationalized through the lens of the stylized model described above. In particular, I focus on the effects of relaxing the assumption that labor productivity realizations are drawn from independent skill distributions.

Specifically, I assume that the marginal occupational ability distributions are still Frechet, as in Subsection 1.4.4, but now their joint distribution is given by $G_\phi(z_l, z_h) = C_\phi\big(G_h(z_h), G_l(z_l)\big)$, where $C_\phi\big(G_h\left(z_h\right), G_l\left(z_l\right)\big) = C_\phi\big(u_h, u_l\big) = -\left(\frac{1}{\phi}\right) \ln\left(1 + \frac{\left(e^{-\phi u_h}-1\right)\left(e^{-\phi u_l}-1\right)}{e^{-\phi}-1}\right)$ is a bi-variate Archimedean Copula. [22] with parameter ϕ [23] [24] [25]

[21] This can be shown by computing the wage premium elasticity of the ratio of average occupational labor income.

[22] See Nelsen (2006).

[23] Informally, Copulas are functions that join or *couple* multivariate distribution functions to their one-dimensional marginal distribution functions. More formally, according to Sklar's Theorem (1959), for every d-dimensional joint distribution function F, with marginals $F_1, \ldots, F_d$, there exists a Copula C such that $F(x_1, ..., x_d) = C(F_1(x_1), ..., F_d(x_d))\ \forall\ x_i \in [-\infty, \infty]$ and $i = 1, ..., d$. Moreover, if the marginals F_i are continuous, the copula $C(\cdot)$ is unique.

[24] Copulas have the advantage that they contain all the information on the dependence structure between variables, whereas the marginal CDFs contain all the information about the marginal distributions. Thus, the properties of the marginal distributions presented in Section 1.4.4 are preserved.

[25] Archimedean copulas have the advantages that they admit explicit formulas and that they allow to model dependence in arbitrarily high dimensions with only one parameter. In what follows I use ϕ to denote

Under these assumptions, average labor income in high- and low-skill occupations are

$$\bar{z_h} = \int_0^{+\infty} z_h \left(\int_0^{z_h\left(\frac{w_h}{w_l}\right)} G_\phi(z_l, z_h) dz_l \right) dz_h,$$

$$\bar{z_l} = \int_0^{+\infty} z_l \left(\int_0^{z_l\left(\frac{w_l}{w_h}\right)} G_\phi(z_l, z_h) dz_h \right) dz_l.$$

where $G_\phi\big(G_h\left(z_h\right), G_l\left(z_l\right)\big)$ is joint cumulative distribution function of (Z_h, Z_l), modeled via a Copula with parameter ϕ. Even though Archimedean Copulas admit explicit formulas, in general these expectations can not be computed in closed form.

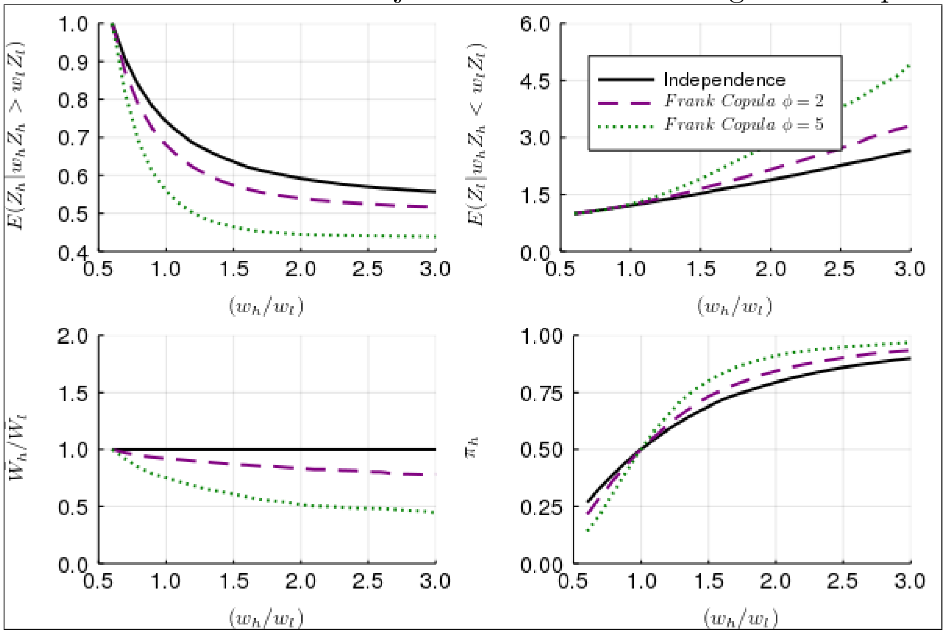
To make clear what the consequences of relaxing the Independence assumption are, I simulate the four main model objects under interest: expected labor productivity in high- $(\mathbb{E}\big(Z_h|w_h z_h > w_l z_l\big))$ and low-skill $(\mathbb{E}\big(Z_l|w_l z_l \geq w_h z_h\big))$ occupations, the fraction of workers in high-skill occupations (π_h), and the ratio of average labor incomes in high-skill with respect to low-skill occupations $(\frac{\bar{W_h}}{\bar{W_l}})$. I experiment with different dependence parameter values and compare the results with the independence case. The results are presented in Figure 1.5 below

The main message from Figure 1.5 is that, once we allow for dependence, the ratio of labor average labor incomes in high- and low-skill occupations is no longer invariant to changes in the skill-premium [26]. The higher the dependence is, the higher the observed decline in average labor productivity in high-skill occupations, the higher the increase we see in average labor income in low-skill occupations, and the higher the corresponding decline in the

the parameter that governs the strength of dependence.

[26]I experimented with negative dependence and it has the opposite results as desired. Similar results can be obtained through lower-tail dependence, using a Clayton Copula. On the other hand, upper-tail dependence modeled trough a Gumbel Copula leads to an increasing average labor income ratio.

Figure 1.5: Simulated Model Objects Under Different Degrees of Dependence



Note: 1,000 simulations of 10,000 draws each are carried out assuming common shape $\theta_l = \theta_h = 2$ and scale parameters $S_l = S_h = 0.50$ for the marginal Frechet densities. Marginal distributions dependence is modeled through a Frank Copula, where $\phi \in \mathbb{R} \setminus \{0\}$ represents the dependence parameter. Average labor productivity in high- and low-skill occupations and the resulting ratio is normalized to one for an initial skill-premium value of 0.5.

average high-to-Low labor income ratio.

The intuition is as follows. Under positive correlation between ability draws, the observed comparative advantage for a given worker is now smaller than what it would be in the independent case. Thus, for a comparable increase in the skill premium more workers are likely to switch from low- to high-skill occupations. In the Frechet case descried in Subsection , since comparative advantage is positively correlated with absolute advantage, workers in high-skill occupations are now expected to have a smaller ability, which leads to a faster decline in average labor productivity in high-skill occupations. The opposite holds for low-skill occupations. As a consequence, the ratio of average labor productivities falls at a higher rate than in the independence case, more than compensating for the increase in the skill-premium, which leads to a decline in the average labor income ratio.

1.4.6 Endogenous Human Capital Accumulation.

I now proceed to relax the assumption that skills are exogenously determined. In particular, I assume that workers can make a discrete and costly decision to build up skills through education. For the sake of tractability, I let individuals choose between no schooling and two different levels of education only ($e = \{s, u\}$), which can be thought of completing Secondary (s) and University (u), respectively. To fix ideas, denote by $\beta_j^e > 0$ the logarithm of the return to completing educational level e in occupation j and c_j^e the corresponding educational fixed cost. For simplicity, assume that the fixed cost of education is common across occupations and only differs across schooling levels ($c_h^s = c_l^s = c^s$ and $c_h^u = c_l^u = c^u$). Additionally, I suppose that the returns to acquiring the lowest educational level are common across occupations ($\beta_h^s = \beta_l^s = \beta^s$), while the yield of completing the highest educational level is higher in high-skill than in low-skill occupations ($\beta_h^u > \beta_l^u$).

Proposition 2 (**Educational Sorting Conditional on Occupational Choice**). *Suppose that the fixed cost of attaining the highest level of education c^u is high enough.*

1. *Conditional on working in high-skill occupations, workers' educational decision is characterized by thresholds $z_h^\star = \frac{c^s}{w_h\left(e^{\beta^s}-1\right)}$ and $z_h^{\star\star} = \frac{c^u-c^s}{w_h\left(e^{\beta_h^u}-e^{\beta^s}\right)}$ such that individuals with $z < z_h^\star$ acquire no education, workers $z_h^\star \leq z_h < z_h^{\star\star}$ complete the first level of education only, while individuals with $z_h \geq z_h^{\star\star}$ attain the highest level of education possible.*

2. *Conditional on working in low-skill occupations, workers' educational choice is defined by thresholds $z_l^\star = \frac{c^s}{w_l\left(e^{\beta^s}-1\right)}$ and $z_l^{\star\star} = \frac{c^u-c^s}{w_l\left(e^{\beta_l^u}-e^{\beta^s}\right)}$ such that individuals with $z_l < z_l^\star$ opt out of education, and workers with $z_l^\star \leq z_l < z_l^{\star\star}$ and $z_l \geq z_l^{\star\star}$ complete the first and second occupational levels, respectively.*

Proof. See Appendix A.5. □

Having fully characterized workers educational decisions conditional on their occupational choices, I proceed to describe the rules that pin down occupational choice.

Proposition 3 (Occupational Choice Under Endogenous Education). *Denote by* $\tilde{z}_l^{\star\star} = z_h^{\star\star}\left(\frac{w_h}{w_l}\right) = \frac{c^u - c^s}{w_l\left(e^{\beta_h^u} - e^{\beta^s}\right)}$. *Under endogenous human capital accumulation through schooling, the occupational selection decisions are as follows:*

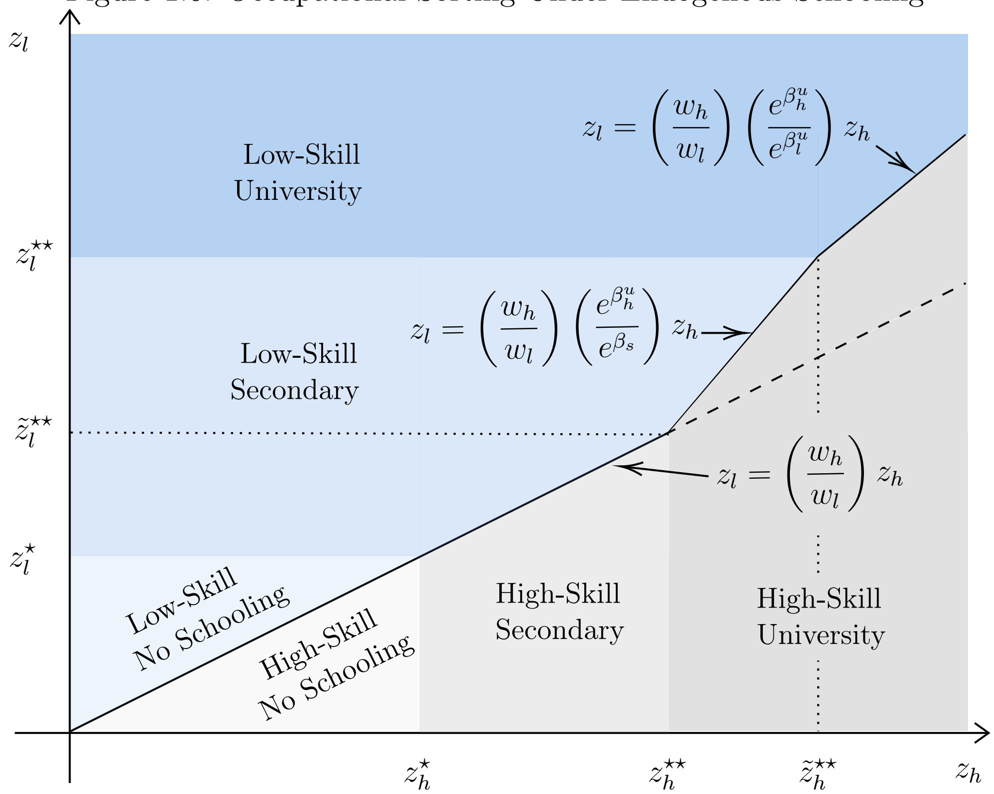
1. *If* $z_l \leq \tilde{z}_l^{\star\star}$ *workers choose high-skill occupations if* $z_h > \left(\frac{w_l}{w_h}\right) z_l$.
2. *If* $z_l \in (\tilde{z}_l^{\star\star}, z_l^{\star\star}]$ *workers choose high-skill occupations if* $z_h > \left(\frac{w_l}{w_h}\right)\left(\frac{z_l e^{\beta^s}}{e^{\beta_h^u}}\right)$.
3. *If* $z_l > z_l^{\star\star}$ *workers choose high-skill occupations if* $z_h > \left(\frac{w_l}{w_h}\right)\left(\frac{z_l e^{\beta_l^u}}{e^{\beta_h^u}}\right)$.

Proof. See Appendix A.5. □

The intuition for Proposition 3 is summarized in Figure 1.6 above, which provides a full description of the occupational and educational choices of individuals, for given efficiency wages w_h and w_l.

Item 1 encompasses several possible cases. In all of them, the occupational decision is based on workers' *raw* comparative advantage. This is because workers either find it optimal not to get education ($z_h < z_h^\star$ and $z_l \leq z_l^\star$), or to get Secondary education in both occupations ($z_h^\star \leq z_h \leq z_h^{\star\star}$ and $z_l^\star < z_l \leq \tilde{z}_l^{\star\star}$), which under common returns does not change their comparative advantage. Additionally, Item 1 includes four more cases where workers educational choice modifies their raw comparative advantage ($z_h^\star < z_h \leq z_h^{\star\star}$ and $z_l \leq z_l^\star$, $z_h \leq z_h^\star$ and $z_l^\star < z_l \leq \tilde{z}_l^{\star\star}$, $z_h \leq z_h^\star$ and $z_l > \tilde{z}_l^{\star\star}$, $z_h^\star < z_h \leq z_h^{\star\star}$ and $z_l > \tilde{z}_l^{\star\star}$). However, in all of these cases individuals' comparative advantage before education is already high enough to

Figure 1.6: Occupational Sorting Under Endogenous Schooling



choose either high- or low-skill occupations, and schooling only augments that comparative advantage in their occupation of choice under the *raw* comparative advantage rule. Hence, comparing workers comparative advantage with relative wages provides a sufficient occupational choice rule for all the cases included in Item 1.

In Item 2 of Proposition 3 the occupational choice decision is modified to take into account that in that region workers go to University if they choose high-skill occupations and, at most, they complete Secondary school if they select low-skill occupations. Thus, since the returns to University are higher than those to going to Secondary school, in this region the schooling decision augments workers' comparative advantage more than proportionately in high-skill occupations. As a consequence, individuals who would not have chosen high-skill occupations according to their *raw* comparative advantage might choose to do so now. This includes the set of individuals between the solid and the dashed line in this region.

Finally, the intuition of Item 3 in Proposition 3 is very close to the one explained above, except for the fact that in this region, if workers were to choose low-skill occupations they would also acquire University education. Since the returns to going to University are higher in high-skill occupations, the comparative advantage of workers grows proportionately more in high-skill occupations, and individuals with who would not have chosen high-skill occupations according to their *raw* comparative advantage might choose to do so now. Again, these workers are the ones whose comparative advantage after education is between the solid and the dashed line in this region.

Proposition 4 (**Endogenous Variables Under Education**). *Denote by* $\tilde{z}_h^{\star\star} = z_l^{\star\star}\left(\frac{w_l}{w_h}\right)\left(\frac{e^{\beta^s}}{e^{\beta_h^u}}\right) = \frac{c^u - c^s}{w_h\left(e^{\beta_l^u} - e^{\beta^s}\right)}\left(\frac{e^{\beta^s}}{e^{\beta_h^u}}\right)$.

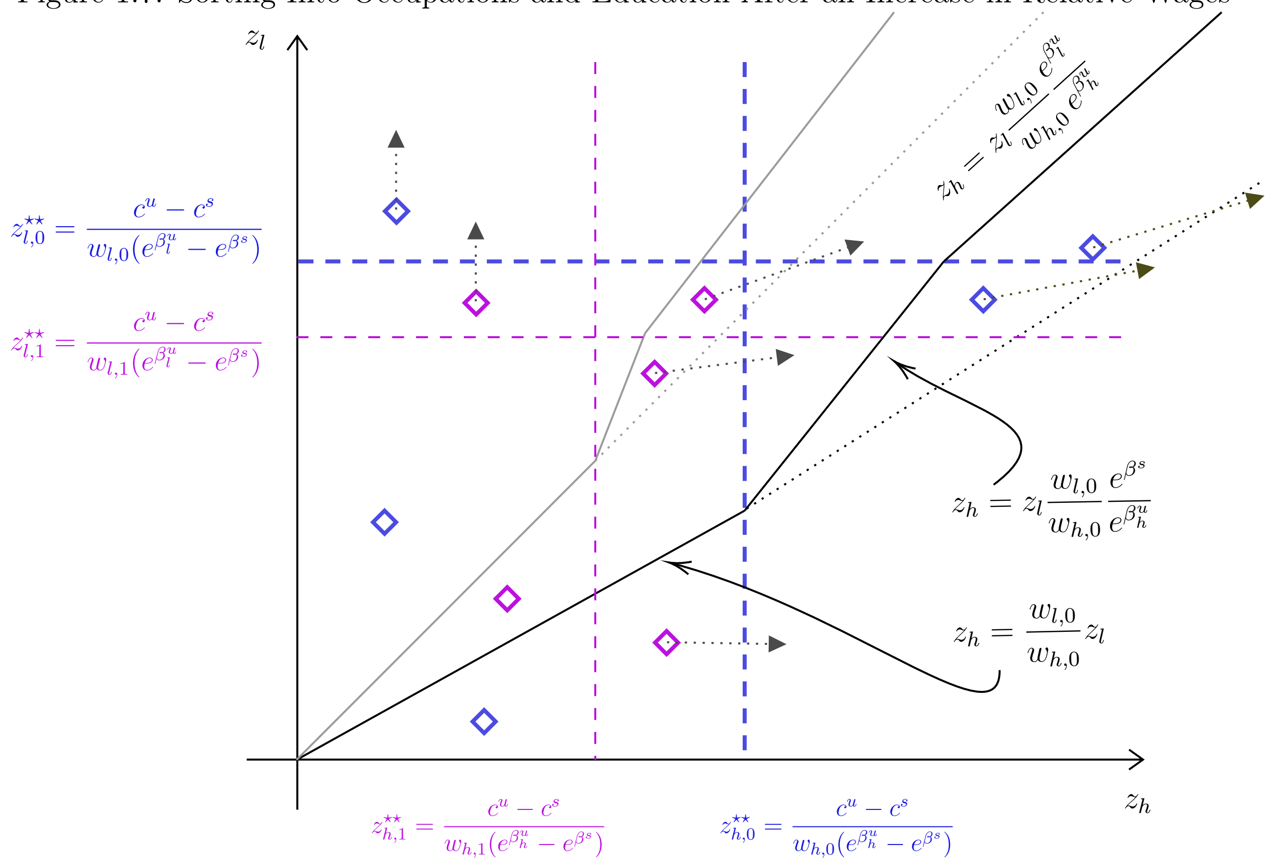
1. *The share of workers in high-skill occupations is given by:*

$$\pi_h = \int_0^{z_h^{\star\star}} \int_0^{\left(\frac{w_h}{w_l}\right) z_h} g_\phi\left(z_h, z_l\right) dz_h dz_l + \int_{z_h^{\star\star}}^{\tilde{z}_h^{\star\star}} \int_0^{\left(\frac{w_h}{w_l}\right)\left(\frac{e^{\beta_h^u}}{e^{\beta^s}}\right) z_h} g_\phi\left(z_h, z_l\right) dz_h dz_l$$
$$+ \int_{\tilde{z}_h^{\star\star}}^{+\infty} \int_0^{\left(\frac{w_h}{w_l}\right)\left(\frac{e^{\beta_h^u}}{e^{\beta_l^u}}\right) z_h} g_\phi\left(z_h, z_l\right) dz_h dz_l \tag{1.9}$$

Proof. See Appendix A.5. □

Proposition 4 is useful to understand how the share of workers in high-skill occupations respond the changes in absolute and relative wages under endogenous education. Assuming that the development process is characterized by skilled-biased changes in productivity that raise both absolute (w_h and w_l) and relative efficiency ($\frac{w_h}{w_l}$) wages, we can use this equation to understand what forces drive the increase in the share of skilled labor as countries develop. These are mainly two channels. The first mechanism is standard and is captured by an increase in the upper limits of integration in the three inner integrals on the right hand side of Equation 1.9: a higher high-to-low skill efficiency ratio attracts lowers the comparative advantage required by workers to choose high-skill occupations, which attracts more individuals to them.

Figure 1.7: Sorting Into Occupations and Education After an Increase in Relative Wages



The second mechanism is the educational channel, which is particularly relevant when increased access to education modifies workers' comparative advantage. This might happen because the relative increase in efficiency wages make it proportionately more profitable to go to university in high-skill occupations, as captured by an increase in the distance of the integration limits of the outer integral ($\uparrow z_h^{\star\star\star} - z_h^{\star\star}$) on the second term on the right hand side of the Equation 1.9, or because more workers now acquire University education in whatever occupation they choose, but the higher returns to University in high-skill occupations lead to an increase in their after education comparative advantage. The latter is captured a reduction in the lower limit of integration on the outer integral in the third term of the right hand side of Equation 1.9 ($\downarrow z_h^{\star\star\star}$).

1.5 Parameter Estimation.

In this section I describe the strategy I follow to discipline the model parameters. I study a quantitative version of the model where the joint skill distribution is fully flexible, allowing for different scale and shape parameters in the marginal distributions and for dependence between marginals. Additionally, I let workers accumulate human capital through education, as in Subsection 1.4.6 above.

The goal is to obtain values for thirteen model parameters. The first three objects are technology parameters: the elasticity of substitution between occupational labor services σ, and the technological efficiency parameters in high- (A_h) and low-skill occupations (A_l). To calibrate the elasticity of substitution between labor types I follow Caselli and Coleman and set σ to 1.42. The relative technological efficiency parameters $\frac{A_h}{A_l}$ are calibrated to match the relative wage per efficiency unit of labor in high-skill occupations with respect to low-skill occupations for the US in 2015, which is obtained as explained in Section 1.4.3, and the fraction of output per hour worked that corresponds to labor for the US in 2015 [27].

[27] Assuming an aggregate technology of the form $y = k^{\alpha}\left[(A_l L_l)^{\left(\frac{\sigma-1}{\sigma}\right)} + (A_h L_h)^{\left(\frac{\sigma-1}{\sigma}\right)}\right]^{\left(\frac{\sigma(1-\alpha)}{\sigma-1}\right)}$ this requires information on GDP, number of people employed, average hours worked per employee and the capital stock for the US in 2010, which I obtain from Penn World Table 9.2. To calibrate the capital share parameter I follow Caselli and Coleman and set $\alpha = 1/3$.

The remaining ten parameters are related to the joint distribution of skills: the shape and scale parameters of the high-skill innate ability marginal distribution θ_h and S_h, the shape and scale parameters of the low-skill innate ability marginal distribution θ_l and S_l, the parameter that governs the dependence between marginals in the Frank Copula, ϕ, the parameters that represent the fixed cost of acquiring University c^u and Secondary education c^s, and the log-returns to Secondary β_l^u and University β_h^u, β_l^u education.

Leaving aside the log-returns to education for now, the skill distribution and educational cost parameters do not have direct data counterparts and depend on how workers' select themselves into occupations and education, so I estimate them via a Simulated Method of Moments. This procedure requires at least seven data moments that I match to seven model-simulated counterparts, and themselves are functions of the deep models parameters under interest. The results of the estimation algorithm is a set of parameters that minimize the weighted sum of the squared distances between the data targets and their model simulated analogs.

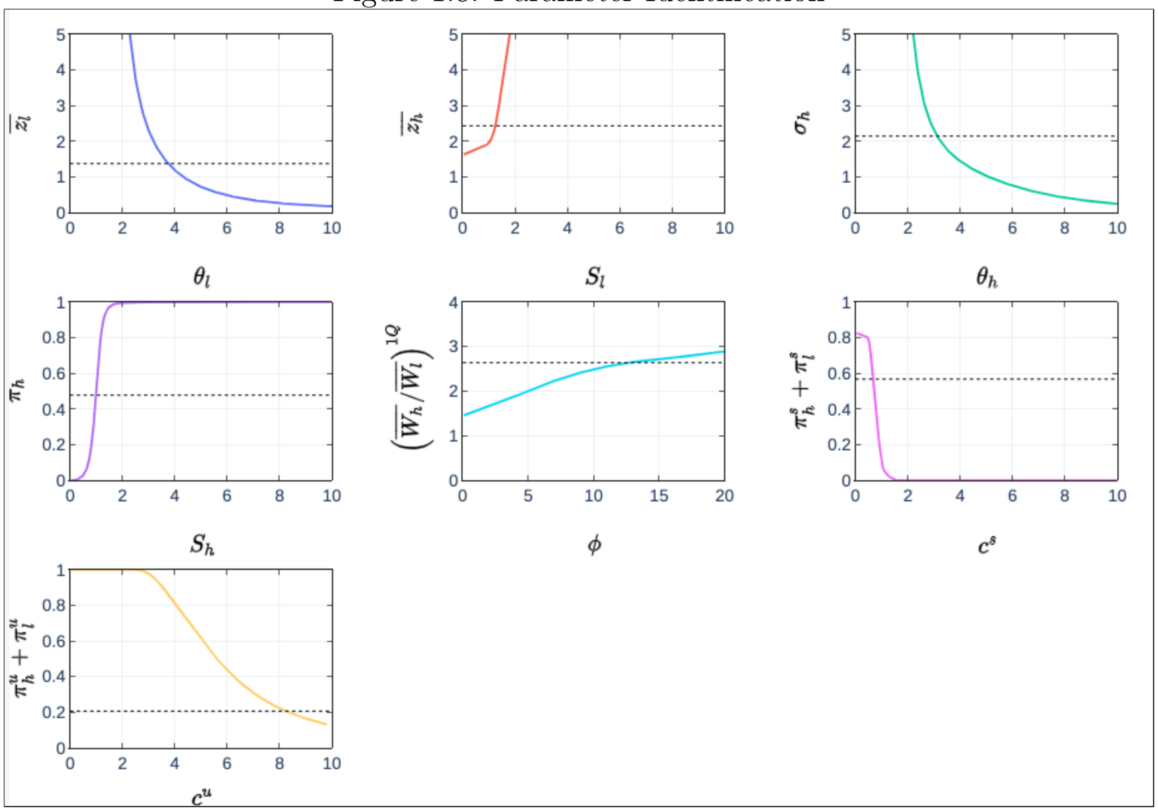
The data moments under interest come from the occupational labor productivity distributions I infer from the data, following the procedure described in Section 1.4.3. The data moment description and their corresponding values are presented in Table 1.3 below.

Table 1.3: Description of Data Moments Targeted

Target	**Description**	**Target Value**
$\mathbb{E}(Z_l \vert w_l Z_l \geq w_h Z_h)$	Mean of the labor productivity distribution in low-skill occupations (US)	1.85
$\sigma_{(Z_h \vert w_h Z_h > w_l Z_l)}$	Standard deviation of the labor productivity distribution in high-skill occupations (US)	2.15
$\mathbb{E}(Z_h \vert w_h Z_h > w_l Z_l)$	Mean of the labor productivity distribution in high-skill occupations (US)	2.44
$(W_h/W_l)^{IQ}$	Avg. Earnings in high- relative to low-skill occupations for the (Avg. IQ country)	2.64
π_h	Fraction of workers in high-skill occupations (US)	0.478
$\pi_h^s + \pi_l^s$	Fraction of workers with Secondary education (US)	0.568
$\pi_h^u + \pi_l^u$	Fraction of workers with University education (US)	0.207

Figure 1.8, on the other hand, shows the model-simulated analogs for these moments and how they react to changes in the underlying parameters. It matches a particular parameter with the model-simulated moment that is more sensitive to it, and plots in the blue lines how the simulated moment varies with changes in the parameter value, fixing the values for the other parameters under interest, and in the red line the targeted value.

Figure 1.8: Parameter Identification



Note: Model-simulated moments are calculated by running, for a given set of parameter values, 1,000 model simulations. The blue line shows how the value of the simulated moment changes as I change the value of one parameter at a time while leaving the remaining parameters fixed at the values that minimize the weighted sum of squared errors in the Simulated Method of Moments. In each figure, the X axis represents a model parameter and the Y axis a moment.

Finally, to estimate the log returns to education I use, once again, U.S. Census data on worker's wage and salary income, main occupation, hours worked, labor force attachment and demographic characteristics from IPUMS International. To be precise, the returns are estimated by regressing log hourly wages for full-time, full year workers by occupational group on three education dummies (less than Secondary, Secondary complete and University dropouts, University graduates and above), a quartic in experience, interactions of the education dummies and the experience quartic, and two race categories (white, other).

Table 1.4 below presents the values for the calibrated and estimated model parameters. It

Table 1.4: Calibrated/Estimated Values for the Model Parameters

Parameter	Parameter Description	Parameter Value
	Calibrated Parameters	
σ	Elasticity of substitution between labor types	1.50
A_h	Technological Efficiency of high-skill occupations	3.38
A_l	Technological Efficiency of low-skill occupations	0.87
	Parameters Estimated Through SMM	
θ_h	Shape parameter of high-skill labor productivity distribution	2.83
		(0.5084)
S_h	Scale parameter of high-skill labor productivity distribution	0.99
		(0.0010)
θ_l	Shape parameter of low-skill labor productivity distribution	3.72
		(0.0028)
S_l	Scale parameter of low-skill labor productivity distribution	1.31
		(0.0081)
ϕ	Dependence parameter in the Frank Copula	12.11
		(0.5220)
c^s	Fixed cost of Secondary education	0.67
		(0.0041)
c^u	Fixed cost of University education	8.19
		(0.1483)
	Log-returns to education	
β^s	Average log return to Secondary school	0.12
		(0.0010)
β^u_h	Log return to University in high-skill occupations	0.59
		(0.0027)
β^u_l	Log return to University in low-skill occupations	0.44
		(0.0011)

1.6 Quantitative Analysis.

I begin assessing the model's prediction for relative labor income in high-skill occupations with respect to low-skill occupations. To that end, for each level of GDP per capita I target the fraction of workers in high-skill occupations (π_h) and the share of individuals that acquire Secondary and University education ($\pi^s = \pi_h^s + \pi_l^s$ and $\pi^u = \pi_h^u + \pi_l^u$). This three targets are enough to pin down relative efficiency wages ($\frac{w_h}{w_l}$) and the fixed costs of education (c^s and c^u), which suffice to fully characterize the occupational and educational selection decisions of workers' at each level of GDP per capita. Average labor productivity in high- and low-skill occupations can be pinned down as a result.

One more comment is in order. The procedure described above requires associating different levels of GDP per capita with their corresponding share of workers in high-skill occupations and the share of workers that acquire each level of education. To obtain a smooth relationship, I run, for the countries in my sample, an OLS regression on each of these two variables on a quadratic for GDP per capita. The results are summarized in Figure 1.9 below.

As below above, the model reproduces the empirical pattern of the relative labor income ratio in high-skill versus low-skill occupations. In particular, it mimics the negative relationship documented in Section 1.1. In absolute terms, the model's relative labor income prediction is below the actual values we observe in the data, particularly for the 4th development quartile. To be precise the actual average labor income ratios in high-skill versus low-skill occupations are 2.6 in the first development quartile, declining to 2.4, 1.8, and 1.6 in the second, third, and fourth quartiles, respectively. In the case of the model's prediction, these values are 2.8,2.3,2.1, and 2.0. Overall, the model does fairly well in terms of reproducing the empirical behavior of labor income. I take this as evidence that workers' selection

Figure 1.9: Relative Hourly Labor Income in the Model vs the Data

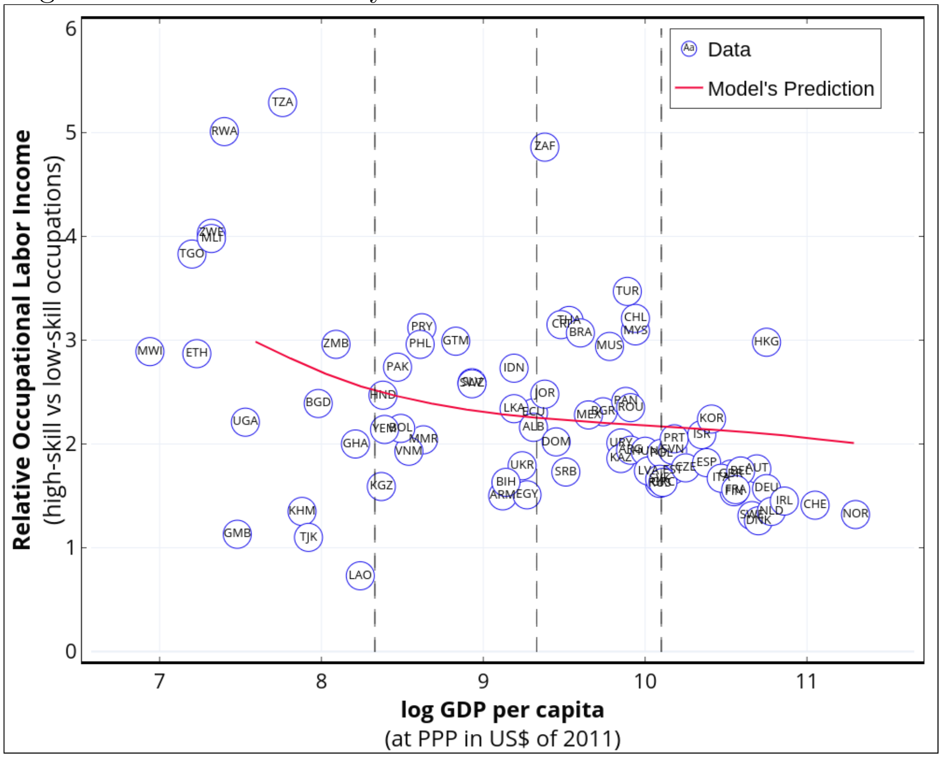
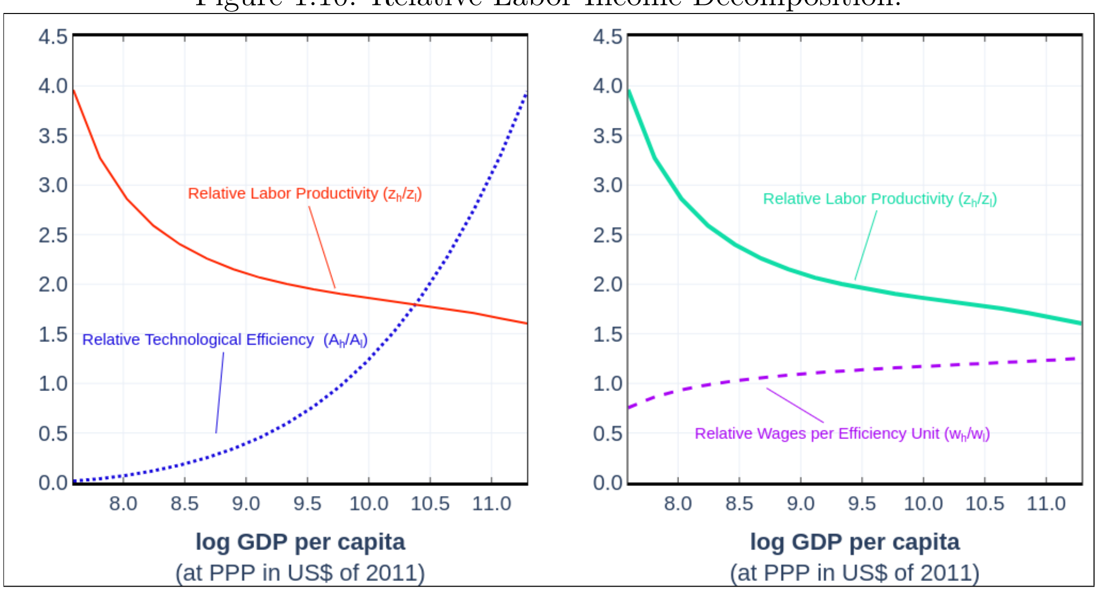


Figure 1.10: Relative Labor Income Decomposition.



into jobs due to observable and unobservable skills is an important feature to understand the relationship between skills and technological progress across development.

The left panel of 1.10 displays how average abilities in high-skill relative to low-skill occupations and the relative technological efficiency of high-skill labor evolve with development, according to the model. As shown in figure, the average ability of workers in high-skill occupations with respect to those in low-skill jobs falls as countries grow richer. This is particularly more pronounced in the first development quartile. At the same time, relative technological efficiency in high-skill occupations with respect to its low-skill occupations counterpart grows monotonically with GDP per-capita, driving the development process.

The right panel of Figure 1.10 decomposes the model's relative average labor income prediction into a relative average ability $\frac{\bar{z_h}}{\bar{z_l}}$ and a relative wages per efficiency unit component $\frac{w_h}{w_l}$. The main message of this figure is that the predicted decline in relative earnings in high- versus low-skill occupations is as a consequence of a decline in the ratio of average ability in high-to-low-skill occupations that more than compensates the rise in relative wages per efficiency unit of labor as countries grow richer.

1.6.1 Effective Skills versus Technologies.

In here use the quantitative version of the model to decompose the decline in labor income in high-skill versus low-skill occupations across development. In particular, I focus on the observed decline between the average country in the first development quartile and the US. To do so, I fix the skill distribution parameters at the values estimated for the US and target, for each country, the fraction of workers in high-skill occupations and the share of educated individuals. In order to match these targets I adjust three model parameters: the ratio of technological efficiency in high-skill with respect to low-skill occupations ($\frac{A_h}{A_l}$), and the fixed costs that workers need to incur to complete Secondary and University schooling (c^s and c^u).

To calibrate these parameters and pin down the equilibrium corresponding to them in each country, I use the results presented in Proposition 4. To be precise, I use bisection methods to first find the relative efficiency wage ($\frac{w_h}{w_l}$) and ability thresholds ($z_h^{\star\star}$ and $z_l^{\star\star}$) that rationalize workers' occupational and higher-level educational decisions. Since the decision to get Secondary education does not affect occupational choices, with these three endogenous variables in hand I can compute the Secondary schooling ability thresholds ($z_h^{\star}$ and $z_l^{\star}$) that reconcile the observed educational attainment levels.

These three endogenous variables are sufficient to fully characterize the occupational and educational choices of workers. Thus, with them one can obtain the effective labor supply in high- and low-skill occupations. To find the values of the calibrated parameters of interest, I back out $\frac{A_h}{A_l}$ from

$$\frac{w_h}{w_l} = \left(\frac{\pi_h \bar{z}_h}{\pi_l \bar{z}_l}\right)^{-\left(\frac{1}{\sigma}\right)} \left(\frac{A_h}{A_l}\right)^{\left(\frac{\sigma-1}{\sigma}\right)}.$$

Finally, the absolute values of c^u and c^s are obtained by solving the system of two equations in two unknowns defined by $\frac{z_h^{\star\star}}{z_h^{\star}}$ and $\frac{z_l^{\star\star}}{z_l^{\star}}$ define.

The calibrated parameters and equilibrium model variables needed to perform the labor income decomposition are presented in Table 1.5 below.

Table 1.5: Relative Labor Income Decomposition in the Model

	$\frac{w_h}{w_l}$	c^u	c^s	$\frac{A_h}{A_l}$	A_h	A_l	$\frac{\bar{z}_h}{\bar{z}_l}$	$\frac{\pi_h}{\pi_l}$	$\frac{L_h}{L_l} = \frac{\pi_h \bar{z}_h}{\pi_l \bar{z}_l}$
US	1.26	8.19	0.67	3.89	3.38	0.87	1.50	0.90	1.35
Avg IQ Country	0.92	3.74	0.41	0.055	0.018	0.33	2.85	0.09	0.26

The counterfactual values for the labor income ratio are exhibited in Table 1.6 below. The left panel of presents the labor income ratio for the US in the first row. The second and third rows show the counterfactual relative labor income for a hypothetical country with the US's relative technological efficiency and the average first development quartile's country raw (l.ii) and effective relative labor supply (l.iii). In the fourth row (l.iv) I let the relative technological efficiency adjust to the level calibrated for the average country in the first development quartile. As a consequence, the fourth row presents the level of relative occupational earnings corresponding to the average country corresponding to the poorest 25 per cent of my sample.

By virtue of the calibration procedure, the model matches the actual decline in earnings in high-skill occupations with respect to low-skill occupations as countries grow from the first development quartile to the level of development observed in the US. In what follows, I decompose the fraction of the decline in labor income explained by the model into a relative technological efficiency in high-skill occupations and a relative quality of labor in high- versus low-skill occupations.

As row (l.ii) shows, in a counterfactual country with the US's technological efficiency and the average first development quartile's country raw supply of skills, the labor income ratio rises to 8.76. Moreover, letting relative labor productivity adjust to the average first development quartile's value increases the relative labor income ratio to even further to 10.86. Thus, the goal is to explain a total change in the relative labor income ratio of 8.22 (10.86-2.64). Since the increase in the relative quality of skilled labor leads to a change in relative labor income of 2.10 (10.86-8.76), it follows that 25% of the change in relative labor income ratio is due to relative skilled labor productivity, while the remaining 75% is due to differences in the relative efficiency of the technology that uses skilled labor in production.

A similar exercise can be performed by adjusting relative technological efficiency first and computing the effect of relative labor productivity residually. In the context of a linear model, these two exercises would give the same answer. It turns out that in this case the

Table 1.6: Relative Labor Income and Counterfactuals

	$\bar{W_h}\backslash\bar{W_l}$			$\bar{W_h}\backslash\bar{W_l}$	
(l.i)	US	1.91	(r.i)	US	1.91
(l.ii)	US $\frac{A_h}{A_l}$ and $\frac{\bar{z_h}}{\bar{z_l}}$ - IQ $\frac{\pi_h}{\pi_l}$	8.76	(r.ii)	US $\frac{\bar{z_h}}{\bar{z_l}}$ and $\frac{\pi_h}{\pi_l}$ - IQ $\frac{A_h}{A_l}$	0.46
(l.iii)	US $\frac{A_h}{A_l}$ - IQ $\frac{\pi_h}{\pi_l}$ and IQ $\frac{\bar{z_h}}{\bar{z_l}}$	10.86	(r.iii)	US $\frac{\bar{z_h}}{\bar{z_l}}$ - IQ $\frac{A_h}{A_l}$ and IQ $\frac{\pi_h}{\pi_l}$	2.11
(l.iv)	Avg IQ Country	2.64	(r.iv)	Avg IQ Country	2.64

two answers differ by a small magnitude. Row two in the right panel of Table 1.6 shows that fixing the effective supply of skills at US's level but adjusting the relative technological efficiency to the average first development quartile's country would have led to a decline in relative labor income to 0.46. Thus, the goal is to decompose a total change of 2.18 (2.64-0.46) in the relative income ratio. Letting the relative raw supply of high-skilled labor adjust to the level in less-developed countries leads to an increase in the income ratio to 2.11. It follows that the residual increase from 2.11 to 2.64 (0.53) is due to differences in the relative high-skill labor productivity. This is roughly a 25% (0.53/2.11).

1.6.2 Measured Gaps in Cross-Country Skilled Labor Efficiency.

Table 1.5 above presents the values of the relative ($\frac{A_h}{A_l}$) and absolute (A_h and A_l) technological efficiency parameters for the US and the average country in the less-developed group. As we can see, the model implies that there exist sizeable gaps in high-skill labor technological efficiency. In fact, the skill-bias in technologies is about 70 times higher in rich than in poor countries, and mostly driven by vast gaps in the absolute efficiency of the technologies that employ high-skill labor. For instance, A_h is 190 times higher in the US than in the average country in the first development quartile, while A_l is only about 2.6 times higher in rich countries.

Based on data on educational attainment and returns to education, Rossi finds that relative efficiency is, at least, 100 times higher in rich than in poor countries. However, the poorest country in his sample is Indonesia, which in my case, belongs to the second development quartile. To contrast my model predictions with those in 1.5 I redo my exercise for Indonesia. The results are presented in Table 1.7 below.

Table 1.7: Relative Labor Income Decomposition for the US and Indonesia

	$\frac{w_h}{w_l}$	c^u	c^s	$\frac{A_h}{A_l}$	A_h	A_l	$\frac{\bar{z}_h}{\bar{z}_l}$	$\frac{\pi_h}{\pi_l}$	$\frac{L_h}{L_l} = \frac{\pi_h \bar{z}_h}{\pi_l \bar{z}_l}$
US	1.26	8.19	0.67	3.89	3.38	0.87	1.50	0.90	1.35
Indonesia	0.97	4.72	0.48	0.09	0.063	0.59	2.63	0.12	0.32

As we can see, measured relative technological efficiency is, roughly speaking, 54 times higher in my case, which is around half of what Rossi finds.

In another recent study, using a method that that relies on disaggregated trade and industry data, Malmberg finds relative skilled labor efficiency to be between 3 and 28 higher rich and poor countries. As in the case of Rossi, the poorest countries in his sample belong to the second development quartile in my dataset. Under his preferred calibration, the relative efficiency of skilled labor is 6.5 times higher in rich countries. However, one major difference with Malmberg is that in his work he focuses on the manufacturing sector, while I study aggregate gaps in skilled labor efficiency. This can translate into major quantitative differences, as gaps in technological efficiency are the lowest in Manufacturing and Services and the highest in Agriculture [28].

1.6.3 Discussion: Absolute and Relative Skill Bias in Technology.

In addition to relative technological efficiency ($\frac{A_h}{A_l}$), Table 1.5 presents the absolute values for the technology parameters in high- and low-skill occupations (A_h and A_l). As we can see,

[28] See Rossi and Gollin et al.

the absolute technological efficiency in both high- and low-skill occupations improves with development, with a much more pronounced in the case of the high-occupations. In fact the absolute efficiency of the high-skill technology is roughly 60 times higher in the US than in the average country in the less-developed group. On the other hand, the absolute efficiency parameter in the low-skill labor technology is only about 2.5 times higher.

These results differ from those found by Caselli and Coleman (CC), who find that the absolute efficiency of unskilled labor falls with development. Notice that the difference with CC is due to my finding that average labor quality of skilled labor is higher in poor countries. In an alternative hypothetical country with the same characteristics as the average nation in the first development quartile of my sample and the same relative quality of skilled labor as in the US, the skill-bias in technology for the poor country would be 0.011, or about five times smaller that in the benchmark case (0.055).

1.6.4 The Effects of an Educational Expansion

I continue to explore the model's quantitative implications of an educational expansion. Within the model, this counterfactual exercise can be performed by analyzing the effects of a reduction in the fixed costs of education c^s and c^u. The most interesting case from policy perspective is studying if increased access to education can push forward less-developed countries towards higher standards of living. Hence, I analyze the effects of a reduction in both educational costs for the average country in the first development quartile. The engineered reduction in educational costs is such that least-developed countries reach the same level of educational attainment as those observed in the group of wealthiest nations after the thought policy is in place.

To do so, I first calibrate the model to match the employment rate in high-skill occupations and the percentage of workers with Secondary and University education in the first development quartile, as explained in Section 1.6.1 above. With the model calibrated to match the most salient labor market and aggregate outcomes for the first development quartile, I proceed to perform a reduction in the costs of acquiring Secondary and University education (c^s and c^u) until the educational attainment levels in the average poor country in my sample are the same as those observed in the US. Since to perform this exercise I keep relative and absolute technological efficiency fixed at the pre-expansion levels, finding the new equilibrium requires pinning down the level of wages per efficiency unit of labor (w_h,w_l) that clear the labor market, which together with the educational costs are enough to fully characterize workers' occupational and educational decision. The results of the exercise are summarized in Table 1.8 below.

A relatively cheaper access to schooling more than doubles the percentage of workers that complete Secondary education (56.9% vs 15.8%). Roughly speaking, this is entirely driven by increased Secondary school completion by individuals in low-skill occupations, which rises from 15.8% to 56.9%. The fraction of workers with University education rises more than proportionately (20.7% vs 3.6%), but in this case, driven by a rise in completion by individuals in high-skill occupations. Since the educational attainment rules are monotonically increasing in ability, the 46.2 percentage point increase in the fraction of workers with University complete in high-skill occupations is given in part by workers who previously would have only completed Secondary in 26.4 percentage points and by individuals who would not have acquired education otherwise in 19.8 percentage points. Something similar occurs in low-skill jobs, but in this case 15.7 percent of workers who were only finished Secondary school before the policy now go to University, while the increased access to Secondary school comes from individuals that would not have accessed education before the expansion (45.3 percent).

Moving on to the occupational structure, more affordable access to education reduces the

Table 1.8: Effects of an Educational Expansion

Variable	Baseline	Educational Expansion ($\downarrow c^s, c^u$)	Abs/% Change
π_h	8.5%	5.2%	-3.3 p.p.
$\pi_h^s + \pi_l^s$	15.8%	56.9%	41.1 p.p.
$\pi_h^s \backslash \pi_h$	28.8%	2.4%	-26.4 p.p.
$\pi_l^s \backslash \pi_l$	14.6%	59.9%	45.5 p.p.
$\pi_h^u + \pi_l^u$	3.6%	20.7%	17.1 p.p.
$\pi_h^u \backslash \pi_h$	33.4%	79.6%	46.2 p.p.
$\pi_l^u \backslash \pi_l$	0.8%	16.5%	15.7 p.p.
w_h	1.18	0.80	-32.2%
w_l	1.28	1.17	-8.6%
$\bar{z}_h$	4.74	7.04	48.5%
$\bar{z}_l$	1.66	1.97	18.7%
$\bar{W}_h \backslash \bar{W}_l$	2.63	2.43	-7.6%
Output	1.000	1.15	15.0%
c^u	3.74	1.40	-58.8%
c^s	0.41	0.18	-56.1%
$c^u \backslash$Avg. Earnings	1.5	0.6	-
$c^s \backslash$Avg. Earnings	0.3	0.1	-

fraction of the population who works in high-skill occupations. However, when compared with the effects on the educational structure, this reduction is marginal. The slight fall in the equilibrium share of workers in high-skill occupations (-3.3 p.p.) is driven by general equilibrium effects, which come from the fact that the effective supply of labor increases more than proportionately in high-skill occupations, reducing relative efficiency wages in high-skill versus low-skill occupations.

Continuing with the results, even though the occupational structure is almost unchanged, the educational expansion has a positive effect on output per-worker, which grows 15 percent, as a consequence of increased average labor productivity, mainly in high-skill jobs.

Compared to a model where workers are classified into high- and low-skill according to their educational attainment, the growth effects of the educational expansion are between one-third and one-tenth model, depending on if the educational attainment threshold to be considered high-skill is Primary school complete or University complete. In these types of frameworks, since labor quality is purely determined by educational attainment, as larger fractions of the population access education the effective labor supply increases. In the case of skilled labor, both the raw and the effective quantity of skilled labor increases. On the other hand, the effective supply of unskilled labor depends on if the quality effect coming from higher educational attainment within a group compensates or not the fall in the quantity of workers below any given threshold. For the Primary complete case, the quantity effect dominates, while in the University case, the quality effect offsets the quantity effect. Thus, the growth effects are larger in the latter case.

The results are in line with the effects of a major occupational expansion that happened in Brazil between 1990 and 2010. As Jaume (2019) documents, Brazil implemented several important educational reforms to increase the educational level of the population. These reforms included, among others: an increase in public expenditure on education from 2.0 percent of GDP in 1995 to over 5.0 percent in 2010 and reducing the direct and indirect cost of schooling by creating more schools and universities together with conditional cash transfers, which I consider to be empirical counterpart of a reduction of the fixed cost of education (c_e) in my model.

Interestingly, Jaume finds that the share of workers with secondary education doubled from 20.5 to 40.0 percent, while the share with university grew from 11.3 to 23.6 percent, and the share of workers with only primary education or less halved from 68.1 to 36.4 percent. In addition, the occupational structure of employment remained relatively fixed, with workers of all educational groups increasingly employed in lower wage occupations. For example, there was an increase in employment of only 1.9 percentage points in the one third of occupations with the highest wages in the economy, despite the expansion of 13 percentage points in the share of high educated workers. These results are in line with the predictions of my model.

However, he finds that inequality measures improved in Brazil, mostly driven by the fact that wages for primarily educated workers soared as a consequence of a reduced supply and an increase in demand for these type of workers. My model predictions in terms of inequality are ambiguous and depend on the group under consideration.

1.7 Conclusion.

I study how relative efficiency and the relative quality of skilled and unskilled labor vary with development. Using harmonized, occupational labor market outcomes for a broad set of countries across the development spectrum, I document that employment in high-skill occupations, or jobs that are relatively more intensive in non-routine cognitive tasks, grows with development. In addition, workers earnings in high-skill occupations falls with respect to those in low-skill occupations as countries grow richer, with elasticities in line with those found by studies based on educational attainment and Mincerian returns to education.

To shed light on these findings and disentangle the mechanisms that determine the relative quality and efficiency of skilled labor, I build a general equilibrium model of occupational choice and human capital accumulation through education. In the model, exogenous skill-biased shifts in productivity attract workers with a lower comparative advantage to occupations that are more skill intensive, in the sense that their abstract task component is relatively higher. The resulting average labor productivity of workers in high- and low-skill occupations depends on how their comparative and absolute advantage is correlated, which depends on the properties of the joint distribution of skills in the population.

I discipline the joint skill distribution and other model features using US labor market data and use the quantitative version of the model to study how the relative efficiency and the relative quality of skilled and unskilled labor vary with development. I find that in poor countries, the relatively small share of workers in high-skill occupations is composed by individuals that are both relatively and absolutely more productive in performing them. In addition, there is positive selection of workers into education, with a higher fraction of those in high-skill occupations achieving higher levels of educational attainment. As a consequence, the relative quality of skilled labor is higher in poor than in rich countries.

When used to decompose the decline in relative labor income of high-skill and low-skill occupations between poor and rich countries, I find that relative quality explains between a quarter and a third of this fall, while relative efficiency explains the remaining 70-75 percent. Additionally, the fact that the relative productivity of workers in high-skill occupations is higher in less develop countries doubles the measured gap in relative efficiency between rich and poor countries.

Chapter 2

Skill-Biased Structural Change

Francisco J. Buera Joseph P. Kaboski Richard Rogerson

Juan Ignacio Vizcaino [1]

2.1 Introduction

The dramatic increase in the wages of high skilled workers relative to low skilled workers is one of the most prominent secular trends in the US and other advanced economies in recent decades. Isolating the underlying causes of this trend is important for projecting future trends and evaluating the extent to which policies might be effective or advisable. The literature has consistently concluded that skill-biased technological change (SBTC) is a quantitatively important driver of the increase in the relative demand for high skilled workers.[2] In this paper we argue that a distinct process – which we label skill-biased *structural*

[1]Buera: Department of Economics, Washington University in St. Louis, One Brookings Drive, St. Louis, MO 63130 (email: *fjbuera@wustl.edu*); Kaboski: Department of Economics, University of Notre Dame, 3060 Jenkins Nanovic Hall, Notre Dame, IN 46556 (email: *jkaboski@nd.edu*); Rogerson: Department of Economics, Princeton University, Julis Romo Rabinowitz Building, Princeton, NJ 08544 (email: *rdr@princeton.edu*); Vizcaino: Department of Economics, Washington University in St. Louis, One Brookings Drive, St. Louis, MO 63130 (email: *jivizcaino@wustl.edu*)

[2]Important early contributions to the literature on the skill premium that stress skill biased technical change include Katz and Murphy (1992b), Bound and Johnson (1992), Murphy and Welch (1992), Berman

change – has also played a quantitatively important role. We use the term skill-biased structural change to describe the *systematic* reallocation of sectoral value-added shares toward high-skill intensive industries that accompanies the process of continued development among advanced economies.

The economic intuition behind our finding is simple. If (as we show is indeed the case in the next section) the process of development is systematically associated with a shift in the composition of value added toward sectors that are intensive in high-skill workers, then the demand for high-skilled workers will increase, independently of whether development is driven by skill-neutral or skill-biased technical change. This channel is absent in analyses that adopt an aggregate production function, since in that case development that comes from skill-neutral technical change has no effect on the relative demand for high-skilled workers.

To assess the quantitative significance of this channel, we develop a simple general equilibrium model of structural transformation that incorporates an important role for skill and use it to study the evolution of the US economy between 1977 and 2005. In order to best highlight the shift in value added to high skill-intensive sectors, we study a two-sector model in which the two sectors are distinguished by their intensity of high-skill workers in production. We allow for sector-specific technological change, which is a (sector-specific) combination of skill-neutral and skill-biased technical change. We show how the model can be used to infer preference parameters and the process for technical change using data on the change in the composition of employment by skill, the change in aggregate output, changes in sectoral factor shares, the skill premium, relative sectoral prices and the distribution of sectoral value added.

In the data, our measure of the skill premium increases from 1.41 to 1.90 between 1977

et al. (1994) and Berman et al. (1998). This is not to say that SBTC is the *only* factor at work, as the literature has also highlighted the effect of other factors on overall wage inequality. For example, DiNardo et al. (1996) argue that labor market institutions such as minimum wages and unionization have played an important role in shaping wage inequality overall, Feenstra and Hanson (1999) emphasize the role of offshoring, and Autor et al. (2013) emphasize the role of trade.

and 2005, an increase of 49 percentage points.[3] Our calibrated model perfectly matches this increase. We then use the model to decompose this increase into three different components: one due to the changes in the relative supply of high-skill workers, one part that is due to skill-biased technical change, and a third part due to other technological changes. If there had been no change in technology, our model predicts that the increase in the relative supply of high-skill workers would have lowered the skill premium to 0.88, a drop of 53 percentage points. It follows that overall changes in technology created an increase in the skill premium of 102 percentage points. In our benchmark specification, between one quarter and one third of this increase comes from changes in technology other than skill-biased technical change, operating through their effect on the composition of value added. We conclude that systematic changes in the composition of value added associated with the process of development are an important factor in accounting for the rise in the skill premium. In fact, if skill-biased technical change had been the sole source of technical change over this period, our model predicts that the skill premium would have increased by only 22 percentage points instead of by 49.

Having established the importance of this effect for the US, we repeat the analysis for a set of nine other OECD countries. While there is some variation in the contribution of compositional changes in value added to changes in the skill premium across countries, ranging from around 15 percent to slightly less than 50 percent, the average for this sample is 25 percent, very much in line with our estimates for the US.

Our paper is related to many others in two large and distinct literatures, one on SBTC and the skill premium and the other on structural transformation. Important early contributions to the literature on the skill premium include Katz and Murphy (1992b), Bound and Johnson (1992), Murphy and Welch (1992), Berman et al. (1994) and Berman et al. (1998). Given

[3]Our measure of the skill premium compares those with *at least* college degrees to those with high school degrees *or less*, and is based on total compensation and not just wages and salaries, which explains why this increase is larger than what the literature typically reports.

that the increase in the skill premium occured in the face of a large increase in the supply of high skill workers, all of these papers sought to identify factors that would increase the relative demand for high-skilled workers. In addition to skill-biased technological change, each of them noted compositional changes in demand as a potentially important element of the increased relative demand for skill. Relative to them, our contribution is fourfold. First, we document the importance of compositional effects that are *systematically* related to the process of development. Second, we show how to uncover the different dimensions of technological change in a multi-sector framework. Third, we present a general equilibrium model in which one can assess the driving forces behind compositional changes. Fourth, and perhaps most importantly, our structural approach finds a much larger role for compositional effects.

An early contribution in the second literature is Baumol (1967), with more recent contributions by Kongsamut et al. (2001) and Ngai and Pissarides (2007). (See Herrendorf et al. (2014) for a recent overview.) Relative to this literature our main contribution is to introduce heterogeneity in worker skill levels into the analysis and to organize industries by skill intensity rather than broad sectors.

The paper that we are most closely related to is Buera and Kaboski (2012). Like us, they study the interaction between development and the demand for skill, though their primary contribution is conceptual, building a somewhat abstract model to illustrate the mechanism. Relative to them our main contribution is to build a simple model that can easily be connected to the data and to use the model to quantitatively assess the mechanism. Leonardi (forthcoming) considers a similar mechanism to us, but focuses on how demand varies by education attainment as opposed to income more broadly, and finds relatively small demand effects.[4] An important antecedent of our work is the paper by Acemoglu and Guerrieri

[4] Ngai and Petrongolo (2014) use a similar framework to show that compositional changes in value added associated with development can explain part of the decrease in the gender wage gap that has occured in the US over time.

(2008). Like us, they study the relationship between development and structural change in a model that features heterogeneity in factor intensities across sectors. But differently than us, they focus on differential intensities for physical capital and the role of the relative price of physical capital rather than human capital. Their work is also primarily theoretical.

An outline of the paper follows. Section 2 presents aggregate evidence on the relation between development and the value added share for high skill intensive services in a panel of advanced economies, in addition to some other important empirical patterns. Section 3 presents our general equilibrium model and characterizes the equilibrium. Section 4 shows how the model can be used to account for the evolution of the US economy over the period 1977 to 2005, and in particular how the data can be used to infer preference parameters and the process of technical change. Section 5 presents our main results about the contribution of various factors to the evolution of the skill premium. Section 6 assesses the contribution of skill-biased structural change for relative prices, and in Section 7 we extend our analysis to a set of nine other countries. Section 8 concludes.

2.2 Empirical Motivation

This section documents the prominence of what we refer to as skill-biased structural change, as well as some of its salient features. In particular, using data for a broad panel of advanced economies, we document two key facts. First, there is a strong positive correlation between the level of development in an economy, as measured by GDP per capita, and the share of value added that is attributed to high skill services. Second, there is also a strong positive correlation between the level of development and the price of high skill services relative to other goods and services. Interestingly, these relationships are very stable across countries, and in particular, the experience of the US is very similar to the average pattern found in the data.

We supplement the above aggregate time series evidence for a panel of countries with some evidence about cross-sectional expenditure shares in the US economy. In particular, we show that the expenditure of higher income households contains a higher share of high skill intensive value-added. This fact will serve two purposes. First, it is suggestive evidence for a non-homotheticity in the demand for high skill services, which is a feature we will include in our model. Second, this cross-sectional moment provides important information about preference parameters that is not readily available from aggregate time series data.

2.2.1 Aggregate Panel Evidence

The starting point for our analysis is the earlier work of Buera and Kaboski (2012). They divide industries in the service sector into two mutually exclusive groups: a high skill intensive group and a low skill intensive group, and show that whereas the value added share of the high skill intensive group rose substantially between 1950 and 2000, the value added share of the low skill intensive group actually fell over the same time period. This finding suggests that the traditional breakdown of economic activity in the structural transformation literature, into agriculture, manufacturing and services, is perhaps not well suited to studying the reallocation of economic activity in today's advanced economies. Here we pursue this line of work further, modifying their aggregation procedure to include goods-producing industries, and extending their analysis to a broad panel of advanced economies.

The analysis is based on value added data from the EUKLEMS Database ("Basic Table").[5] These data exist in comparable form for a panel of 12 advanced economies over the years 1970-2005.[6] The sectoral value-added data are available at roughly the 1 to 2-digit industry level. We focus on a two-way split of industries into high skill intensive and low skill intensive

[5]See O'Mahony and Timmer (2009).

[6]These countries are Australia, Austria, Denmark, France, Germany, Italy, Japan, the Netherlands, South Korea, Spain, the United Kingdom, and the United States. The U.S. data for value-added go back to only 1977, while the Japan data go back to only 1973.

based on the share of labor income paid to high-skill workers.[7] While one could imagine more detailed splits, including more than two skill categories and perhaps interacting skill intensity with goods vs. services, we feel that this two-way split both facilitates exposition and allows us to capture a robust pattern in the cross-country data.

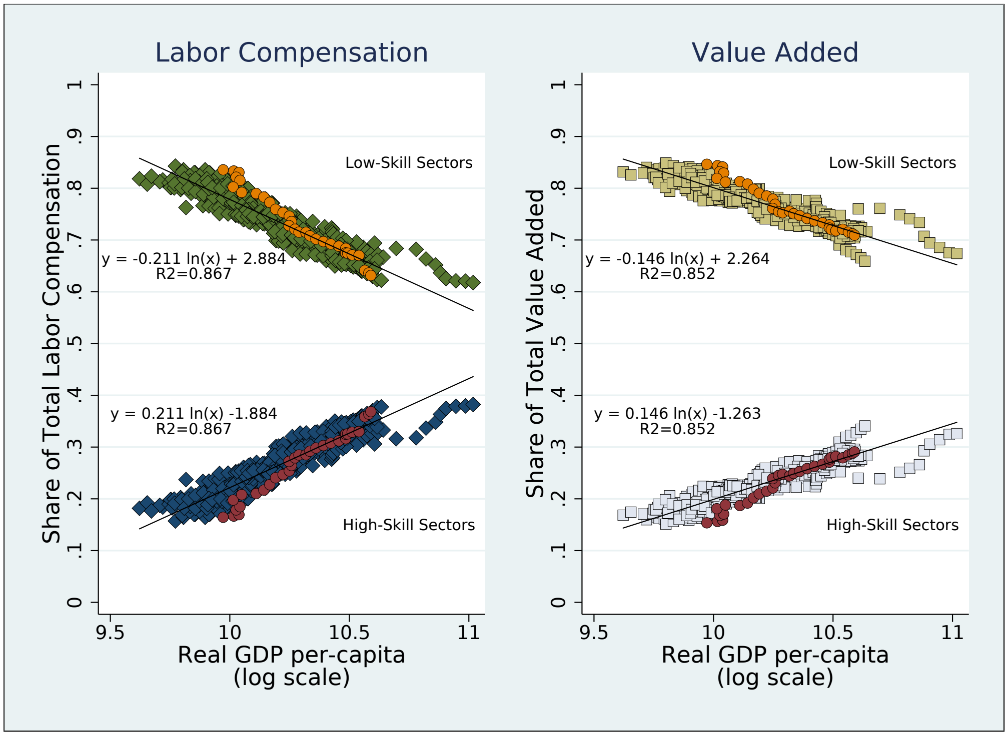
The labor payment data come from the EUKLEMS Labour Input Data and are slightly more disaggregated. High skill-intensive service sectors are: “Financial Intermediation”, “Renting of Machinery and Equipment and Other Business Activities”, “Education”, and “Health and Social Work”. In 1970, the economy-wide average share of labor compensation paid to high-skill workers in the U.S. was 20 percent; the corresponding shares for these high skill-intensive industries were 34, 38, 74, and 49 percent, respectively. These industries remain well above average throughout the time period.[8] We combine these data with real (chain-weighted) GDP per capita data from the Penn World Tables 7.1. Finally, we demean both the value-added share data and the (log) GDP per capita data by taking out country fixed effects.

Figure 1 shows the data pooled across time and countries. The small squares show the relationship for the panel of advanced countries; we have highlighted the United States data using the larger circles. The relationship is clear: the value added share of the high skill-intensive sector increases with log GDP/capita, with a highly significant (at a 0.1 percent level) semi-elasticity of 0.17. The regression line implies an increase of roughly 24 percentage points as we move from a GDP per capita of $10,000$ to $40,000$ (in 2005 PPP terms), and explains 80 percent of the variation in the data. Moreover, we see that the United States data is quite similar to the overall relationship. Indeed, the tight relationship suggests that cross-country differences in the details for funding of education or health, for example, are second order relative to the income per capita relationship in terms of their effects on time

[7]High-skill is defined as a college graduate and above.

[8]The next highest industries are “Chemicals and Chemical Products” (27 percent), “Coke, Refined Petroleum, and Nuclear Fuel” (21 percent), and “Electrical and Optical Equipment” (21 percent).

Figure 2.1: Value-added by Skill-intensity and Economic Development.

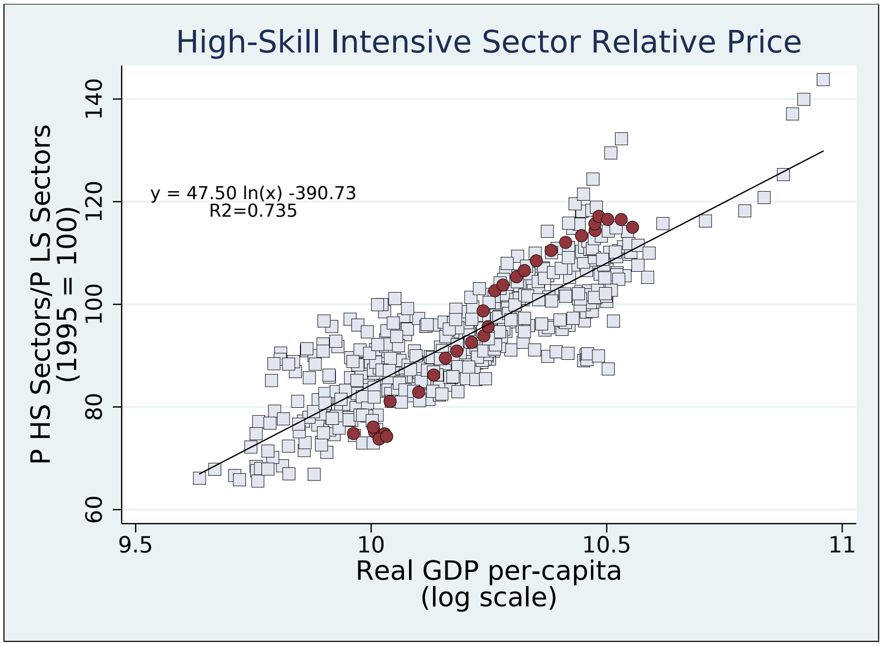


series changes. In sum, the tendency for economic activity to move toward high skill-intensive services as an economy develops is a robust pattern in the cross-country data.

One of the common explanations for structural change is changes in relative prices. (See, for example, Baumol (1967) and Ngai and Pissarides (2007).) Using value-added price indices from the same EUKLEMS Database, we can examine the correlation between the relative price of high-skill intensive services and the increasing value added share of high-skill intensive services that accompanies the process of development.[9] Figure 2 is analogous to Figure 1, but it plots the price index of the high skill-intensive sector relative to the low-skill intensive sector rather than share data on the y-axis. Again we have demeaned both the relative price and log GDP per capita data to eliminate country fixed effects, and normalized the relative price indices to 100 in 1995. As before, the larger circles represent the U.S. data.

[9]We construct sector-level aggregate indices as chain-weighted Fisher price indices of the price indices for individual industries. Calculation details are available in the online data appendix, http://www3.nd.edu/~jkaboski/SBSC_DataAppendix.zip.

Figure 2.2: Relative Price of Skill-intensity Sector and Economic Development



Again, the relationship is striking. The linear regression is highly significant, explains 84 percent of the variation in the demeaned data, and is quantitatively important: the relative price of the high skill-intensive sector increases almost two and a half times over the range of the data. Finally, the U.S. relationship is quite similar to the overall relationship, and again the tight relationship suggests that cross-country variation in this relative price-income relationship is second order. We conclude that changes in relative prices are another robust feature of the structural transformation process involving the movement of activity toward the high-skill intensive sector.

2.2.2 Income Effects: Cross-Sectional Household Evidence

A second common explanation for structural change is income effects associated with non-homothetic preferences. (See, for example, Kongsamut et al. (2001).) With this in mind it is of interest to ask whether high-skill intensive services are a luxury good, i.e., have an income elasticity that exceeds one. To pursue this we examine whether the relationship between the skill intensity of value-added consumption and income exists in the expenditure data from a cross-section of households. To the extent that all households face the same prices at a given point in time and have common preferences (or at least preferences that are not directly correlated with income), the cross-sectional expenditure patterns within a country abstract from the relative price relationship in Figure 2 and allow us to focus on the effect of income holding prices constant.

One complication with pursuing this approach is that it involves mapping household expenditure data through the input-output system in order to determine the consumption shares of value added. We briefly sketch the steps of this procedure here, and provide more details in the online appendix. We start with the household level Consumer Expenditure Survey (CEX) data for the United States from 2012. We adapt a Bureau of Labor Statistics mapping from disaggregated CEX categories to 76 NIPA Personal Consumption Expenditure (PCE) categories and then then utilize a Bureau of Economic Analysis (BEA) mapping of these 76 PCE categories to 69 input-output industries that properly attributes the components going to distribution margins (disaggregated transportation, retail, and wholesale categories). Using the 2012 BEA input-output matrices, we can then infer the quantity of value added of each industry embodied in the CEX expenditures. We classify the 69 industries as high skill intensive or low skill intensive using the EUKLEMS data as previously noted.[10]

This procedure generates household-level data for the share of total expenditure that represents valued added by high skill intensive sectors and low skill intensive sectors, which we can regress on household observables, most importantly income or education, and potentially a host of other household level controls. In our empirical work we restrict ourselves to the primary interview sample, and each observation is a household-quarter observation.

[10]The labour data from EUKLEMS contains 41 distinct industries. The "basic" data, from which we obtain value-added data, contain only 33 distinct industries.

Table 1 presents results for regressions of the total share of expenditures that is high skill intensive. The first column presents results from an OLS regression on log after tax income and a set of demographic controls, including age, age squared, dummies for sex, race, state, urban, and month, and values capturing household composition (number of boys aged 2-16, number of girls aged 2-16, number of men over 16, number of women over 16 years, and number of children less than 2 years). The coefficient on log income in the first column indicates that the semi-elasticity of the share of value-added embodied in expenditures is 0.012. The second column replace log income with the log of total expenditures, yielding a larger semi-elasticity of 0.022.

Table 2.1: Household High-Skill Intensive Expenditure Share vs. Income or Total Expenditures

	OLS	**OLS**	**IV**	**IV**	**OLS**
Ln Income	0.012***	-	0.047***	-	-
SE	0.001	-	0.002	-	-
Ln Expenditures	-	0.022***	-	0.073***	-
SE	-	0.001	-	0.003	-
High Skill Head	-	-	-	-	0.042***
SE	-	-	-	-	0.002
R^2	0.20	0.19	0.05	0.03	0.23
Observations	24,213	27,318	24,213	27,318	8,883

*** indicate significance at the 1 percent level.
Controls include: age; age squared; dummies for sex, race, state, urban, and month; number of boys (2-16 year); number of girls (2-16 years); number of men (over 16 years); number of women (over 16 years);and number of infants (less than 2 years). High skilled is defined as 16 years of schooling attained, while low skilled is defined as 12 years attained.

Income and expenditures are certainly mis-measured in the micro data, and even if properly measured it would only proxy for permanent income, leading to a likely attenuation bias. The second and third columns attempt to alleviate this measurement error by instrumenting for log income or log expenditures using the years of schooling attained by the head of household. Instrumenting for income in this fashion increases the coefficient roughly four-fold to 0.047. Likewise, instrumenting for log total expenditures more than triples the coefficient to

0.073.[11]

The last column use education as a direct regressor, replacing log income or log expenditures with a dummy for whether the head of household is high skilled or not. Here high skill is defined as having exactly 16 years of education, while low skill is defined as having exactly 12 years. (The rest of the households are dropped, leading to the smaller sample.) The coefficient indicates that the share of value-added embodied in expenditures is 4.3 percentage points higher in households with a high-skilled head.

We have examined the robustness of the results in Table 1 along various dimensions. Table 1 uses "quarterly" expenditures of the household across the three months they are surveyed, but if we use the monthly data directly, we find nearly identical results. By defining high skill as those with at least 16 years of education, and low skill as those with less than 16 years of education, we expand the sample somewhat, but the raw estimates are similar (0.032 rather than 0.043). Dropping demographic controls increases the sample by about 15 percent and lowers the coefficients on income by roughly 25 percent and the coefficient on expenditures by roughly one-third, but the coefficients remain highly significant. Dropping the controls has essentially no impact on the high-skilled head of household coefficient. The main effect of dropping the controls is substantially lower R^2 values.

Quantitatively, even the larger, instrumented, expenditure coefficient of 0.073 is substantially smaller than the aggregate time series value of 0.17 in Figure 1, but not negligible in comparison. We therefore take this as suggestive evidence that, in addition to relative prices, non-homotheticicities may also play a role in accounting for the observed pattern of skill-biased structural change.

Lastly, we note an important limitation in directly applying the micro elasticity as an income effect. Because the CEX captures only out-of-pocket expenditures, it underesti-

[11] The larger coefficient for expenditures is driven by certain lumpy expenditures like higher educational expenses and car purchases driving both up in particular months. We nonetheless report these coefficients for the sake of completeness.

mates the true consumption of certain goods like insurance premiums (a substantial share of which is paid by employers) and higher education (a substantial share of which is paid by government).[12]

2.2.3 Summary

In summary, we have documented a robust relationship in the time series data for advanced economies regarding the movement of activity into high-skill intensive services and the process of development. We refer to this process as skill-biased structural change, so as to emphasize both its connection to the traditional characterization of structural change and the special role of skill intensity. This relationship is remarkably stable across advanced economies, thus suggesting that it is explained by some economic forces that are robustly associated with development, with country specific tax and financing systems not playing a central role in explaining the time series changes.

In documenting this relationship we have used a two-way split into high and low skill intensive sectors. This masks important within sector heterogeneity. Indeed, within the low skill intensive sector, a pattern emerges that the relatively more skill intensive sectors within this category, e.g., manufacturing industries like electrical equipment and chemicals, expand relative to the less skill intensive sectors like agriculture or textiles.[13] In this sense, our simple dichotomy may understate the true extent of skill biased structural change. However, the relative price patterns, use patterns (consumption and investment), and trade patterns make the analysis at a more disaggregated level more difficult to interpret and much less

[12] The estimated income semi-elasticity of the share of out-of-pocket insurance is actually significantly negative in the CEX data although overall insurance consumption is certainly positive. Similarly, although the expenditure share-income semi-elasticity of higher education is positive, it is likely understated. Finally, the lack of primary and tertiary expenditures may actually be overstated in the CEX data because it neglects public expenditures, but we conjecture that this relationship is small relative to the higher education relationship.

[13] Katz and Murphy (1992b) give a detailed analysis across 150 2-digit industry-occupation cells for the period, 1963-1987. Autor and Dorn (2013) present a recent account focusing on detailed occupation categories.

directly tied to traditional structural change forces.

The traditional structural transformation literature emphasizes the role of both nonhomotheticities and relative price changes as drivers of structural change, and we have also presented evidence that both of these effects are relevant in the context of skill-biased structural change as well. Specifically, we documented a strong positive correlation between the relative price of high skill intensive services and GDP per capita in a cross-country panel as well as a positive correlation between household value added expenditure shares on high skill intensive services and income in the US cross-section. These relationships are not only highly statistically significant, but they are also economically significant in a quantitative sense.

2.3 Model and Equilibrium

Our analysis emphasizes how intratemporal equilibrium allocations are affected by changes in the economic environment that operate through changes in income and relative prices. To capture these interactions in the simplest possible setting, we adopt a static closed economy model with labor as the only factor of production. Our model is essentially a two-sector version of a standard structural transformation model extended to allow for two labor inputs that are distinguished by skill. In this section, we describe the economy and its equilibrium at a point in time; later we describe the features that we will allow to change over time to generate skill-biased structural change as described in the previous section.

2.3.1 Model

There is a unit measure of households. A fraction f_L are low-skill, and a fraction f_H are high skill, where $f_L + f_H = 1$. All households have identical preferences defined over two commodities. In our empirical analysis these two commodities will be connected to the low and high skill intensive aggregates studied in the previous section. As a practical matter, all of our high skill intensive sectors are services and all goods sectors are in the low skill intensive sector. It is notationally convenient to label the two commodities as goods and services even though what we label as goods includes low-skill services.

We assume preferences take the form: $a_G c_{Gi}^{\frac{\varepsilon-1}{\varepsilon}} + (1-a_G)(c_{Si}+\bar{c}_S)^{\frac{\varepsilon-1}{\varepsilon}}$

where c_{Gi} and c_{Si} are consumption of goods and services by an individual of skill level i, $0 < a_G < 1$, $\bar{c}_S \geq 0$ and $\varepsilon > 0$. Note that if $\bar{c}_S > 0$, preferences are non-homothetic and, holding prices constant, the expenditure share on services will be increasing in income.[14] This non-homotheticity is motivated by the cross-sectional analysis in the previous section. Note that households are assumed to not value leisure, since our focus will be on the relative prices of labor given observed supplies.

Each of the two production sectors has a constant returns to scale production function that uses low- and high-skilled labor as inputs. We assume that each of these production functions is CES: $Y_j = A_j\left[\alpha_j H_j^{\frac{\rho-1}{\rho}} + (1-\alpha_j) L_j^{\frac{\rho-1}{\rho}}\right]^{\frac{\rho}{\rho-1}} j = G, S$

where L_j and H_j are inputs of low- and high-skilled labor in sector j, respectively. The parameter α_j will dictate the importance of low- versus high-skilled labor in each sector. While one could imagine that the elasticity of substitution between these two factors also differs across sectors, our benchmark specification will assume that this value is the same for both sectors. We consider the effects of cross-sectional variation in this parameter in our sensitivity analysis.

Before proceeding to analyze the equilibrium for our model we want to comment on the

[14] This is a simple and common way to create differential income effects across the two consumption categories. One can also generate non-homothetic demands in other ways. For example, Hall and Jones (2007) generate an income elasticity for medical spending that exceeds unity through the implied demand for longevity. Boppart (2014), Swiecki (2014) and Comin et al. (2015) all consider more general preferences with the common feature being that income effects associated with non-homotheticities do not vanish asymptotically. This property is likely to be relevant when considering a sample with countries at very different stages of development. Because we focus on a sample of predominantly rich countries, we have chosen to work with the simpler preference structure in order to facilitate transparency of the economic forces at work.

significance of abstracting from capital and trade. By excluding capital we implicitly adopt a somewhat reduced form view of skill-biased technological change. For example, changes in relative demand for skilled labor due to capital-skill complementarity and changes in the price of equipment (as in (Krusell et al., 2000), for example), will show up in our model as skill-biased technological change. While it is obviously of interest to understand the underlying mechanics of skill-biased technological change, for our purposes we believe our results are strengthened by adopting a more expansive notion of skill-biased technological change rather than focusing on a specific mechanism.

Although we abstract from trade, we view our analysis as complementary to those that emphasize the potential role of trade in shaping the evolution of the skill premium. In particular, our analysis focuses on the extent to which compositional changes between goods and high-skill services diminish the role of skill-biased technological change in accounting for changes in the skill premium. To the extent that trade is dominated by trade in goods, it could diminish the role of skill-biased technical change by potentially affecting compositions within the goods sector. For example, if the US increasingly exported high skill-intensive manufactured goods and imported low skill-intensive manufactured goods, this would shift the composition of production within the goods sector, and in our analysis will be interpreted as skill-biased technological change within the goods sector. Put somewhat differently, we believe that trade may serve to generate a process of skill biased structural change within the goods sector, and in this sense represents an additional channel to the one that we focus on.

2.3.2 Equilibrium

We focus on a competitive equilibrium for the above economy. The competitive equilibrium will feature four markets: two factor markets (low- and high-skilled labor) and two output markets (goods and services), with prices denoted as w_L, w_H, p_G and p_S. We will later normalize the price of low-skilled labor to unity so that the price of high-skilled labor will also represent the skill premium, though in our derivations it will be convenient to postpone implementing this normalization.

The definition of competitive equilibrium for this model is completely standard and straightforward, so here we focus on characterizing the equilibrium. Individuals of skill $i = L, H$ solve

$$\max_{c_{Gi}, c_{Si}} a_G c_{Gi}^{\frac{\varepsilon-1}{\varepsilon}} + (1 - a_G)(c_{Si} + \bar{c}_S)^{\frac{\varepsilon-1}{\varepsilon}}$$

subject to

$$p_G c_{Gi} + p_S c_{Si} = w_i. \tag{2.1}$$

Using the first order conditions of this problem and normalizing w_L to unity, the aggregate expenditure share for services is:

$$\frac{p_S\left[(1-f_H)c_{SL} + f_H c_{SH}\right]}{1 - f_H + f_H w_H} = \frac{1}{\left(\frac{1-a_G}{a_G}\right)^{\varepsilon} + \left(\frac{p_G}{p_S}\right)^{1-\epsilon}} \left[\left(\frac{1-a_G}{a_G}\right)^{\varepsilon} - \frac{p_S \bar{c}_S \left(\frac{p_G}{p_S}\right)^{1-\epsilon}}{1 - f_H + f_H w_H}\right] \tag{2.2}$$

This expression serves to illustrate the two forces driving structural change in the model: relative prices and income effects. Provided that $\epsilon < 1$, as p_G/p_S declines, the expenditure share of services increases. And, provided that $\bar{c}_S > 0$, an increase in income measured in units of services, (i.e., $(1 - f_H + f_H w_H)/p_S$) also leads to an increase in the expenditure share of services.

The problem of the firm in sector $j = G, S$ is

$$\max_{H_j, L_j} p_j A_j \left[\alpha_j H_j^{\frac{\rho-1}{\rho}} + (1-\alpha_j) L_j^{\frac{\rho-1}{\rho}}\right]^{\frac{\rho}{\rho-1}} - w_H H_j - L_j$$

The first order conditions of the firm's problem imply an equation for the price of sector j output in terms of the skill premium w_H :

$$\hat{p}_j(w_H) = \frac{1}{A_j}\left[\frac{\alpha_j^\rho}{w_H^{\rho-1}} + (1-\alpha_j)^\rho\right]^{\frac{1}{1-\rho}}. \tag{2.3}$$

The above expression implies that the search for equilibrium prices can be reduced to a single dimension: if we know the equilibrium wage rate for high-skilled labor then all of the remaining prices can be determined.

Finally, we derive an expression for the market-clearing condition for high-skilled labor that contains the single price w_H. Using $H_j/L_j = \left(\frac{\alpha_j}{1-\alpha_j}\frac{1}{w_H}\right)^\rho$, the production function of sector j, and (2.3), we obtain a sector-specific demand function for high-skilled labor:

$$H_j = \left[\frac{\alpha_j \hat{p}_j(w_H) A_j}{w_H}\right]^\rho \frac{Y_j}{A_j,} \tag{2.4}$$

which, together with equilibrium in the goods market, yields the market-clearing condition for high-skilled labor solely as a function of w_H:

$$\left[\frac{\alpha_S \hat{p}_S(w_H) A_S}{w_H}\right]^\rho \frac{\sum_{i=L,H} f_i \hat{c}_{Si}(w_H)}{A_S} + \left[\frac{\alpha_G \hat{p}_G(w_H) A_G}{w_H}\right]^\rho \frac{\sum_{i=L,H} f_i \hat{c}_{Gi}(w_H)}{A_G} = f_H. \tag{2.5}$$

Here we have used $\hat{c}_{ji}(w_H)$ to denote the demand for output of sector j by a household of skill level i when the high-skilled wage is w_H and prices are given by the functions $\hat{p}_i(w_H)$ defined in (2.3).

2.4 Accounting for Growth and Structural Transformation

In this section we calibrate the model of the previous section so as to be consistent with observations on structural transformation, growth, and the skill premium under the assumption that the driving forces are changes in technology (both skill-biased and skill-neutral) and changes in the relative supply of skill. In particular, we will use the above model to account for observed outcomes at two different points in time, that we denote as 0 and T for the initial and terminal periods respectively. Consistent with the existing literature on technological change and the skill premium, we do not allow the parameter ρ to change over time. We also assume that preferences are constant over time.

2.4.1 Targets for Calibration

Calibrating the model in the initial and terminal period requires assigning values for 14 parameters. Nine of these are technology parameters: 4 values of the α_j (two in each period), 4 values of the A_j (two in each period), and ρ. Three are preference parameters: ε, a_G and $\bar{c}_S$. Lastly we have the value of f_H at the initial and terminal dates. The two initial values of the A_j represent a choice of units, reducing the overall number of parameters to be set to twelve. In our benchmark specification we will set the two elasticity parameters ρ and ε based on existing estimates, further reducing this number to ten.

As described below, we will directly measure the initial and final values of f_H from the data. To calibrate the remaining parameters we will target the following values which reflect the salient features of growth, structural transformation and demand for skill: the initial and final values for factor shares in both sectors, the initial and final value added shares for the two sectors, the initial and final value of the skill premium, the change in the relative price of the two sectors, and the overall growth rate of the economy.

In choosing values for these targets we rely on the EUKLEMS data from Section 2. For the U.S., complete data are available for the years 1977 to 2005, so we choose these two years as our initial and terminal year respectively.[15] This period is of interest, since 1977 effectively marks a local minimum in the skill premium (see (Acemoglu and Autor, 2011b) for earlier data), and it secularly increases after 1977.

Many of the values for targets have direct counterparts in the data and so require no discussion, but the construction of targets for the labor variables does merit some discussion. The data contain total compensation and total hours by industry, skill level ("low", "medium", and "high", which are effectively, less than secondary completion, secondary completion but less than tertiary completion, and four year college degree or more), gender, and age groupings (15-29, 30-49, and 50 and over). Consistent with our calculations in Section 2, we combine the compensation of EUKLEMS categories of "low" and "medium" educated workers of all genders and ages into our classification of low-skilled, and "high" educated workers into our classification of high-skilled, in order to calculate labor income shares by skill at both the aggregate and sectoral level. We use the same sectoral classification as in Section 2.

Setting targets for the skill premium and the relative supply of skilled workers requires that we decompose factor payments into price and quantity components. If all workers within each skill type were identical then we could simply use total hours as our measure of quantity, but given the large differences in hourly wage rates among subgroups in each skill type this seems ill-advised. Instead, we assume that each subgroup within a skill type offers a different amount of efficiency units per hour of work. We normalize efficiency units within each skill type by assuming one hour supplied by a high school-educated ("medium") prime-aged (i.e., aged 30-49) male is equal to one efficiency unit of low skill labor and that one hour supplied

[15]BEA data on value added and prices are also available for the period 1977-2007 and line up quite closely with the KLEMS data. The BEA data does not allow consistent aggregation prior to 1977. Data on labor compensation and hours are only available through 2005, which is why we choose 2005 as our terminal date.

by a college-educated ("high") prime-aged (i.e., aged 30-49) male is equal to one efficiency unit of high skill labor.[16] With this choice of units, the skill premium is defined as the ratio of college-educated ("high") to high school-educated ("medium") prime-aged (i.e., aged 30-49) male wages. This premium rises from 1.41 in 1977 to 1.90 in 2005.[17]

Finally, we infer f_H using the identity that the ratio of labor compensation equals the product of the skill premium and the relative quantity of high- to low-skilled labor (f_H and $f_L = 1 - f_H$, respectively). Equivalently, one could compute efficiency units of each skill type by using relative wages within each skill group to infer efficiency units and directly aggregating efficiency units. Note that our implicit assumption is that differences in wages between different low-skilled (high-skilled) demographic groups reflect differences in efficiency units of low-skilled (high-skilled) labor. This procedure implies that high skill labor was 22% of total labor supply in 1977 and rose to 34% in 2005. (For comparison, the fraction of raw hours that were high skill labor increased from 19% to 31% over the same time period.)

Table 2 summarizes the values used for the targets listed above.

Table 2.2: Values Used to Calibrate Technology Parameters

f_{H0}	f_{HT}	w_{H0}	w_{HT}	$\%\Delta\frac{p_S}{p_G}$	$\%\Delta Y$	θ_{G0}	θ_{GT}	θ_{S0}	θ_{ST}	$\frac{C_{S0}}{Y_0}$	$\frac{C_{ST}}{Y_T}$
0.19	0.31	1.41	1.90	58.9	80.8	0.18	0.35	0.54	0.66	0.19	0.32

[16] While one could obviously normalize units by choosing other reference groups, this group seems most natural since its uniformly high rate of participation over time minimizes issues associated with selection.

[17] Comparing earnings of full time workers using CPS data, Figure 1 in Acemoglu and Autor (2011b) indicates values of 1.39 and 1.64 for 1977 and 2005 respectively. Our measure indicates a roughly 17 percentage point greater increase. This difference reflects two factors. First, although our lower skill group includes those with some college education, our high skill group includes those with post-graduate education, for whom wage growth has been dramatic. Autor et al. (2008) find 31 percentage points log wage growth for those with 18+ years of education between 1979-2005 relative to 18.5 for those with 16 or 17 years. Second, our measure is based on total compensation and not just on labor earnings. Pierce (2010) documents that the change in the 90-10 ratio over this time period is more than twenty log points higher when using total compensation than when using CPS wages. His analysis is based on firm level data and so does not allow a breakdown by educational attainment.

2.4.2 The Calibration Procedure

Having described the targets to be used in the calibration and the data used to determine the values of these targets, we now describe the details of the mapping from targets to parameters. We proceed in two steps. The first step shows how the technology parameters are inferred. In the second step, we describe how to infer values for the preference parameters.

We begin with the determination of technological change. First, we show that given a value for ρ, the four values of the α_{jt} are pinned down by sectoral factor income shares and the skill premium, w_{Ht}. To see this, from equations (2.3) and (2.4) note that the share of sector j income going to high skill labor, $\theta_{Hjt} = \frac{w_{Ht}H_{jt}}{\hat{p}_j(w_{Ht})Y_{jt}}$, is

$$\theta_{Hjt} = \frac{\alpha_{jt}^{\rho}}{\alpha_{jt}^{\rho} + (1-\alpha_{jt})^{\rho} w_{Ht}^{\rho-1}}$$

Therefore, given ρ, the skill premium w_{Ht}, and data for θ_{Hjt}, the value of the α_{jt} are given by:

$$\alpha_{jt} = \frac{1}{1 + \frac{1}{w_{Ht}^{\frac{\rho-1}{\rho}}} \left(\frac{1-\theta_{Hjt}}{\theta_{Hjt}}\right)^{\frac{1}{\rho}}}.$$

Next we turn to determining the values of the A_{jt}'s. As noted previously, the two values in period 0 basically reflect a choice of units and so can be normalized. We will normalize A_{S0} to equal one, and given the calibrated values for the α_{j0} and the value of w_{H0}, we choose A_{G0} so as to imply $p_{G0}/p_{S0} = 1$. In this case p_{GT}/p_{ST} can be easily identified with the change in the relative sectoral prices. As is well known in the literature, with identical Cobb-Douglas sectoral technologies, relative sectoral prices are simply the inverse of relative sectoral TFPs, so the change in relative prices would therefore determine the values of the two A_{jT}'s up to a scale factor.[18] This precise result does not apply to our setting because

[18]This same relation holds more generally, and in particular would also apply if the sectoral production functions are CES with identical parameters.

of sectoral heterogeneity in the α_{jt}'s. (The skill premium also plays a role in determining relative prices, which we examine in Section 6.) Nonetheless, there is still a close connection between relative sectoral prices and relative sectoral TFPs (i.e., the A_{jt}). In particular, using equation (2.3) for the two sectors we have:

$$\frac{A_{Gt}}{A_{St}} = \frac{p_{St}}{p_{Gt}} \left[\frac{\frac{\alpha_{Gt}^{\rho}}{w_{Ht}^{\rho-1}} + (1-\alpha_{Gt})^{\rho}}{\frac{\alpha_{St}^{\rho}}{w_{Ht}^{\rho-1}} + (1-\alpha_{St})^{\rho}} \right]^{1/(1-\rho)}. \tag{2.6}$$

The scale factor will of course influence the overall growth rate of the economy between periods 0 and T, so we choose this scale factor to target the aggregate growth rate of output per worker.[19]

To this point, given a value for ρ, we have identified all of the technology parameters. For our benchmark analysis we set $\rho = 1.42$, which corresponds to the value used in Katz and Murphy (1992b), and which is commonly used in the literature. Though this is a commonly used value in the literature, it is worth noting that previous estimates using an aggregate production function do not necessarily apply in our setting. For this reason we will also do sensitivity analysis with regard to ρ over a fairly wide interval, ranging from 0.77 to 2.50. Table 3 below shows the implied values for the technology parameters.

Table 2.3: Calibrated Technology Parameters ($\rho = 1.42$)

α_{G0}	α_{S0}	α_{GT}	α_{ST}	A_{ST}/A_{S0}	A_{GT}/A_{G0}
0.28	0.55	0.44	0.67	1.30	2.33

A few remarks are in order. Not surprising given the way in which we grouped industries

[19] Note that to compute aggregate output at a point in time (and thus also the growth rate in aggregate output) it is necessary to know the sectoral distribution of output. The relations imposed thus far guarantee that maximum profits will be zero in each sector but do not determine the scale of operation. Intuitively, the split of activity across sectors at given prices will be determined by the relative demands of households for the two outputs. Below we will describe how preference parameters can be chosen to match the sectoral distribution of value added at both the initial and final date. At this stage we simply assume this split is the same as in the data.

into the two sectors, the weight on low-skilled labor is greater in the goods sector than in the service sector at both dates. More interesting is that in both sectors technological change has an important component that is skill biased. While the level rise in α is greater for the goods sector than the service sector, the changes are of similar magnitude (16pp and 12pp).

However, overall technological progress is much greater in the goods sector than in the service sector. The TFP term in the goods sector more than doubles between 1977 and 2005, corresponding to an average annual growth rate of 2.97%. In contrast, the growth of the TFP term in the service sector averages only 0.80% per year.

We now turn to the issue of determining the values for the three preference parameters: a_G, $\bar{c}_S$ and ε. While technological change can be inferred without specifying any of the preference parameters, we cannot evaluate some of the counterfactual exercises of interest without knowing how relative demands for the sectoral outputs are affected by changes in prices. As noted above, the calibration of technology parameters used information about sectoral expenditure shares without guaranteeing that observed expenditure shares were consistent with household demands given all of the prices. Requiring that the aggregate expenditure share for goods (or services) is consistent with the observed values in the data for the initial and terminal date would provide two restrictions on the three preference parameters. It follows that we would either need to introduce an additional moment from the data, or perhaps use information from some outside study to determine one of the three preference parameters. As noted earlier, for our benchmark results we will follow the second approach and fix the value of ε, and then use data on aggregate expenditure shares to pin down the values for a_G and $\bar{c}_S$. Our main finding is relatively robust to variation over a large range of values of ε, thereby lessening the need to tightly determine its value. Nonetheless, in Section 5 we will describe how cross-sectional data on expenditure shares could be used as an additional moment and allow us to determine all three preference parameters.

The empirical literature provides estimates of ε that correspond to the categories of "true" goods and "true" services, but not for our definitions of the two sectors that are based purely on skill intensity. The key distinction is that we have grouped low-skill services with goods. However, given that our goods sectors does contain all of the industries that produce goods, while our service sector does consist entirely of service sector industries, it seems reasonable that information about the elasticity of substitution between the true goods and services sectors should be informative about the empirically plausible range of values for ε in our model. Recalling that the objects in our utility function reflect the value-added components of sectoral output, the relevant estimates in the literature would include Herrendorf et al. (2013), Buera and Kaboski (2009), and Swiecki (2014). All of these studies suggest very low degrees of substitutability between true goods and true services. For this reason we consider values for ε in the set $\{0.125, 0.20, 0.50\}$, with $\varepsilon = 0.20$ chosen as our benchmark.[20]

Given a value for ε, equation (2.2) can be used to determine values for a_G and $\bar{c}_S$ if we require that the model match the initial and final sectoral value added shares.

Table 4 shows the values for the preference parameters in the different scenarios.

Table 2.4: Calibrated Preference Parameters

	ε	a_G	$\bar{c}_S$
Benchmark	0.20	0.97	0.09
High ε	0.50	0.61	0.20
Low ε	0.125	0.99	0.08

The qualitative patterns in this table are intuitive. Note that in each case the changes in income, relative prices and the aggregate expenditure shares are the same. As we move from $\varepsilon = 0.20$ to $\varepsilon = 0.125$ we decrease the elasticity of substitution between the two goods, implying a smaller response in relative quantities but a larger response in relative expenditure shares. In order to compensate for this larger effect, we need to decrease the impact of income changes on relative expenditure shares, implying a lower value for $\bar{c}_S$. The lower value for

[20] Comin et al. (2015) redo the exercise in Herrendorf et al. (2013) for a more general class of preferences and find an elasticity of substitution that is somewhat higher, though still less than 0.50, which is our upper range.

$\bar{c}_S$ will in turn lead to a higher expenditure share on services in the initial period, since the non-homotheticity is now less important. Hence, in order to match the expenditure shares for the initial period we need to attach a lower weight, a_G, to consumption of goods. As we move from $\varepsilon = 0.20$ to $\varepsilon = 0.50$ we see the reverse pattern.

2.5 Decomposing Changes in the Skill Premium

Our model is calibrated so as to account for the observed change in the skill premium between 1977 and 2005. In this section we use the calibrated model to perform counterfactuals that allow us to attribute changes in the skill premium to the various exogenous driving forces in the model. Our primary objective is to decompose the effect of changes in technology on the skill premium into a piece due to skill biased technological change and a residual piece that is due to other forms of technological change. The residual piece affects the relative demand for skilled individuals indirectly, through its impact on the relative consumption of services.

Table 5 reports the results of our counterfactual exercises for each of the three specifications that differ with respect to the value of ε. As we will see, the key results are very similar across the three specifications, so to better focus our discussion we will concentrate on the ε=.20 case and later summarize the other cases.

Table 2.5: Decomposing Changes in the Skill Premium
(US 1977-2005)

	$\varepsilon = 0.50$	$\varepsilon = 0.20$	$\varepsilon = 0.125$
w_{H0}	1.41	1.41	1.41
w_{HT}	1.90	1.90	1.90
w_{HT} - changes in f_i only	0.91	0.88	0.88
w_{HT} - changes in f_i and A_j only	1.10	1.09	1.09
w_{HT} - changes in f_i and α_j only	1.64	1.63	1.63

The first two rows of the table report the starting and ending values for the skill premium, which are the same in our model as they are in the data. The rest of the table decomposes this change into several pieces by considering several counterfactual exercises in our model. The first counterfactual assesses the role of "supply" versus "demand" factors. Specifically, the share of labor supply coming from skilled workers increases between 1977 and 2005, and in the absence of any other changes exerts downward pressure on the skill premium. As noted above, focusing on the ε=.20 case for now, the third row of Table 5 shows that if the change in relative supply of skill (i.e., the f_i's) had been the only change between 1977 and 2005 the skill premium would have decreased from 1.41 to 0.88, a 53 percentage point fall. Given that we in fact observe an increase in the skill premium of 49 percentage points, it follows that the overall effect of technological change is to increase the skill premium by 102 percentage points.

Our next goal is to decompose the 102 percentage point increase in the skill premium due to the overall effect of technological change into one part that is due to skill biased technological change (i.e., changes in the α_{jt}'s) and a second part due to other dimensions of technical change (i.e., changes in the A_{jt}'s).

There are two natural calculations that one could perform to assess the contribution of changes in the A_{jt}'s to changes in the skill premium. In both calculations we start from the previous counterfactual in which we change only the supply of skill. In the first calculation we add in the change in the A_{jt}'s and compute the fraction of the overall 102 percentage point increase that they account for. In the second calculation we instead add in the changes in the α_{jt}'s and compute the fraction of the 102 percentage point increase that is not accounted for. In a linear model these two calculations would give the same answer, but to the extent that nonlinearities are present they may differ. It will turn out that the answers do differ, but only to a relatively minor extent.

The final two rows in Table 5 present the results of these two calculations. Specifically, moving from the third row to the fourth row we see that the change in the A_{jt}'s increases the skill premium from 0.88 to 1.09, an increase of 21 percentage points. This represents roughly 20 percent of the overall 102 percentage point increase accounted for all technical change. Moving from the third row to the fifth row, we see that the changes in the α_{jt}'s cause the skill premium to increase from 0.88 to 1.63. The residual is 27 percentage points, which represents approximately 27% of the 102 percentage point increase due to all changes in technology. Based on this we conclude that non-skill biased technical change accounts for between 20 and 27 percent of the overall change in the skill premium due to technical change. Put somewhat differently, according to our calibrated model, if skill biased technical change had been the only force affecting the relative demand for skill then the skill premium would have increased by only 22 percentage points instead over the period 1977 to 2005 instead of increasing by 49 percentage points.

If we redo these calculations for the other two values of ε the answers are similar. For $\varepsilon = 0.50$ the two methods imply that changes in the A_{jt}'s account for 19% and 26% of the overall change in the skill premium due to technical change, whereas for $\varepsilon = 0.125$ the two values are effectively identical to those for the $\varepsilon = 0.20$ case, being equal to 20% and 27%. From this we conclude that our finding of a significant contribution of changes in the A_{jt}'s is robust to a large variation in the value of ε.

2.5.1 Sources of Structural Change

In the introduction we stressed the fact that aggregate production function analyses abstract from compositional changes, and that our main objective was to assess the quantitative importance of the compositional changes that are associated with the process of structural transformation during development. The previous calculations decomposed the overall changes in the skill premium into parts due to skill-biased technological change and skill-neutral technological change. In order to make the connection between this decomposition and compositional changes it is of interest to examine the connection between technological change and changes in sectoral value added shares. Table 6 reports the results for each of the three values of ε.

Table 2.6: Technical Change and Value Added Share of Skill-Intensive Services
(US 1977-2005)

	$\varepsilon = 0.50$	$\varepsilon = 0.20$	$\varepsilon = 0.125$
Model 1977	0.19	0.19	0.19
Model 2005	0.32	0.32	0.32
Model 2005 with fixed A_j	0.17	0.16	0.16
Model 2005 with fixed α_j	0.35	0.34	0.34

The first two rows of the table remind us that the (skill-intensive) service sector grew significantly between 1977 and 2005, increasing its share of value added from 19 percent to 32 percent. Recall that our calibrated model perfectly replicates the change in the data. The last two rows provide two different ways of assessing the role of changes in the A_j's and the α_j's in accounting for this compositional change. The third row reports the service sector value added share that would have resulted if the changes in the change in the α_{jt}'s had been the only source of technological change, whereas the fourth row reports the service sector value added share that would have resulted if the change in the A_{jt}'s had been the only source of technological change. Both calculations lead to the same conclusion: effectively all of the compositional change is accounted for by changes in the sectoral TFPs. It follows that our previous decomposition of changes in the skill premium due to the two different sources of technical change can effectively be interpreted as statements about the importance of structural change.

Non-skill biased technological change in our model still has two distinct dimensions: one which increases the overall level of TFP in the economy and the other of which increases relative TFP in the goods sector. As we noted above, both of these changes tend to reallocate activity from the goods sector to the service sector, thereby indirectly increasing the relative demand for skill. Next we examine the relative magnitude of these two effects.

Note that for given changes in the A_{jt}'s the relative magnitude of these two effects is dictated by the preference parameters ε and $\bar{c}_s$: as ε becomes smaller, relative TFP changes have larger effects, and as $\bar{c}_S$ becomes larger then sector neutral changes in the A_{jt}'s have larger effects. Because our calibration procedure implies that as ε becomes smaller the value of $\bar{c}_S$ decreases, we expect to find that sector neutral change plays a larger role for smaller values of ε.

To evaluate this we consider the counterfactual in which we hold all parameters fixed from the original calibration, allow the f_{it}'s and the α_{jt}'s to change as before, but counterfactually force the A_{jt}'s to grow at the same rate, with this rate chosen so as to yield the same overall change in aggregate output as in the data. When we do this, the implied values of the skill premium are 0.125 respectively. It follows that when $\varepsilon = 0.50$ it is income effects that dominate the overall impact of the A_{jt}'s on the skill premium, whereas for the smaller values of ε the sector biased nature of TFP growth is somewhat more important than the income effect. So while the three different specifications offer very similar decompositions regarding the overall effect of changes in the A_{jt}'s on the skill premium, they have distinct implications for how different type of changes in the A_{jt} lead to changes in the skill premium.

The preceding discussion has focused on the role of technological change in bringing about changes in the composition of final demand, but one may also ask to what extent increases in the supply of skill can act as a driving force behind structural change. To assess our model's predictions for this we compute the change in the expenditure share for services that would result if the change in the relative supply of skill had been the only driving force. The result is that instead of increasing from .29 to .44, the expenditure share for services actually decreases modestly to .27. Intuitively, there are two effects at work. First, the increase in the relative supply of skill serves to decrease the skill premium and hence the relative price of services. With $\varepsilon < 1$ this leads to a decrease in the services expenditure share. Second, the changes in the skill premium and the price of services lead to a change in income measured in units of services. In our calibrated economy this change in income is positive (i.e., the decrease in the price of services dominates the effect of a decrease in the skill premium), leading to an increase in the services expenditure share. As noted above, the net quantitative effect is a modest decrease. The main message is that increases in the supply of skill do not serve to expand the size of the high skill intensive sector.

2.5.2 Using Data to Infer ε

In the results above we considered a range of values for ε rather than trying to use our model to infer a specific value. Since our main message was robust to a wide range of values for ε we do not view this is a particular limitation of the analysis. Nonetheless, in this subsection we describe how the use of a cross-sectional moment on the household side would allow us to also infer a value for ε.

Intuitively, cross-sectional information can provide information about the magnitude of the income effect: assuming that all households have the same preferences and face the same prices at a point in time, cross-sectional heterogeneity in income will allow us to infer the size of the income effect. This can be implemented using the cross-sectional information that we reported in Section 2. In particular, our empirical analysis found that the expenditure share for services is 0.04 higher for high skill households than for low skill households. If we require that our model match this moment in the final time period, the implied values for ε and $\bar{c}_S$ are 0.05 and 0.08 respectively when we assume $\rho = 1.42$. This would correspond to values of ε somewhat below the lower end of the interval that we considered. As a practical matter, it turns out that moving to increasingly smaller values of ε from $\varepsilon = 0.20$ has relatively small effects, as could already be seen in Table 4. We have also repeated this exercise for values of ρ equal to 0.77 and 2.50 and obtained estimates of $\varepsilon = 0.11$ and 0.01 respectively. While we do not report the results for these cases in detail, we note that our main message remains unaffected if we were to adopt these specifications.

Having offered the idea of using cross-sectional data to infer the size of income effects on the demand for what we have labelled goods and services, we think it is important to repeat one important limitation of this approach in the current context. Two of the largest components of high skill services are education and health care, both of which are not well tracked by household expenditure surveys. To the extent that spending on some components of education and health reflect a collective societal choice, it is not clear that cross-sectional data will be useful in detecting how cross-sectional differences in income affect desired consumption.

2.5.3 Sensitivity Exercises

For the results in the previous section we assumed that $\rho = 1.42$, which we noted was a standard value in the literature, and the value implied by the analysis in Katz and Murphy (1992b). However, we also noted that the aggregate analyses that have supported this estimate are not necessarily appropriate in our multi-sector economy. For this reason we also consider a wider range of values for ρ to assess the extent to which the above conclusions are robust to variation in this parameter.

We consider two alternative values of ρ, corresponding to higher and lower elasticities of substitution. Specifically, we consider $\rho = 0.77$ and $\rho = 2.5$. In each case we redo the calibration procedure as before. While the value of ρ does affect the quantitative findings, it leaves our main message largely unchanged. For example, focusing on the case of $\varepsilon = 0.20$ we find that when $\rho = 0.77$, the share of changes in the skill premium due to technical change that are accounted for by changes in the A_{jt} is 23% and 38% from the two methods. When $\rho = 2.50$ the corresponding values are 15% and 18%. We conclude that our main finding of a significant role for changes in demand composition induced by technical change in accounting for changes in the skill premium is robust to considering a wide range of values for ρ, though higher values of this elasticity parameter do lead to modest declines in the estimated role played by demand composition.

Our analysis has assumed that the value of ρ is the same in both sectors. Absent any empirical evidence on the extent of heterogeneity in ρ across sectors, this seemed a natural benchmark. Reshef (2013) suggests that the elasticity of substitution between high and low-skilled workers may be lower in services, for example. It is therefore important to assess whether our results are sensitive to the assumption of ρ being constant across sectors. To do this we redo our exercise for several specifications in which we allow the two values of ρ to vary across sectors, allowing for the ratio ρ_G/ρ_S to be both larger and smaller than one. In all cases we assume that the weighted average of the two elasticities–$(H_G/H)\rho_G + (H_S/H)\rho_S$–is equal to 1.42 when evaluated at the initial factor shares, so that our analysis can be interpreted as assessing the effect of heterogeneity holding the aggregate elasticity of substitution constant. We consider values for ρ_S of 0.77, 0.91, 1.11, and 2.00, and the implied values for ρ_G are 2.23, 2.06, 1.82, and 0.73. Table 7 reports the same statistics as in Table 5, focusing on the case of $\varepsilon = 0.20$.

Table 2.7: The Effect of Sectoral Variation in ρ
(US 1977-2005)

	$\frac{\rho_S}{\rho_G} = 0.35$	$\frac{\rho_S}{\rho_G} = 0.44$	$\frac{\rho_S}{\rho_G} = 0.61$	$\frac{\rho_S}{\rho_G} = 1.00$	$\frac{\rho_S}{\rho_G} = 2.73$
w_{H0}	1.41	1.41	1.41	1.41	1.41
w_{HT}	1.90	1.90	1.90	1.90	1.90
Counterfactual w_{HT} to changes in:					
f_i only	0.99	0.97	0.94	0.88	0.74
f_i and A_j only	1.11	1.11	1.10	1.09	1.08
f_i and α_j only	1.71	1.70	1.68	1.63	1.48

For ease of comparison, the fourth column repeats the results from our benchmark specification. For values of $\rho_S/\rho_G < 1$ the implications are affected very little, and to the extent that a very large value of ρ_S/ρ_G influences the quantitative results, it yields a larger role for the demand side effects that we focus on (between 29% and 36%). Noting that we are considering a very wide range of variation in the relative values of ρ, we conclude that our results are quite robust to variation in ρ across sectors.

Lastly, we consider the extent to which mis-measurement of relative prices might influence our results. Our quantitative analysis utilized information about changes in the relative price of the high skill intensive sector. Between 1977 and 2005 this relative price increased by more than sixty percent. One possible concern is that price inflation in the high skill intensive sector might be upward biased because of the failure to properly account for quality improvements.

Here we report the results of a simple exercise to assess the extent to which our conclusions are affected by this possibility. In particular, consider the case in which the true increase in the relative price of the high skill intensive sector was only half as much as indicated by the official data. This means that real value added in this sector increased by roughly 30% more than indicated by the official data, and aggregate GDP grew by roughly 15 additional percentage points. We set $\rho = 1.42$ and $\varepsilon = 0.20$ and carry out the same calibration procedure as previously. Not surprisingly, given that we are holding ε fixed and decreasing the role of relative price changes, the calibration procedure yields a larger value for $\bar{c}_S$, indicating a larger role for nonhomotheticicities. However, we find that the contribution of demand factors is virtually identical to what we found in our benchmark calculation. So while mismeasurement of relative price changes has implications for relative magnitudes of preference parameters, it has virtually no effect on our assessment of the role of demand factors.

2.5.4 Comparison with Earlier Literature

The increase in the relative demand for high-skilled labor that we attribute to structural change is substantially higher than the overall effects of relative demand shifts found in the earlier literature on the topic. For example, in the overlapping years of our samples, 1979-1987, Katz and Murphy (1992, KM hereafter) attributed 4.6 percentage points of increase in relative demand for high-skilled labor to changes in industrial composition (i.e., their "between industry" analysis, see Table VIII), and they did so using a much more highly disaggregated industrial classification rather than our stark two-sector specification. This contributes only 11 percent of the increase in skill premium over that period, given their estimated elasticity of substitution and the increased supply over that period. Bound and Johnson (1992) estimate a small but slightly *negative* contribution of industrial composition.

In contrast, when we restrict attention to the overlapping years 1979-1987, our simulations attribute between 22 and 27 percent of the increase in the skill premium to the increase in the relative demand of high-skilled labor associated with skill-biased structural change.

Note that our overall results are for a somewhat later period, 1977-2005 vs. 1963-1987, and our data are substantially more aggregated in terms of industry, educational attainment, and experience levels. Nonetheless, by focusing on overlapping years and conducting KM analyses on our data, we can show that the importance of these differences is quantitatively small and that two important factors stand out in driving our substantially larger estimates for the role of changes in the composition of demand. (See online appendix for more details.)

First, KM's decomposition using observed industrial movements is an approximation, which, as they acknowledge, underestimates the true contribution of demand shifts because the skill premium rose during the period of analysis. A rising skill premium disproportionately increases the price of the skill-intensive output (see equation (8)), reducing the movement of resources into that sector relative to what would be observed with perfectly elastic labor supply (i.e., the full shift in demand).[21] Whereas they are able to analytically sign the bias, the added structure of our model enables us to actually quantify this bias. In addition, our analysis uses a global solution of the model, instead of a local approximation. For the years 1979-1987, our model attributes an increase in the skill premium of between 9 and 11 percentage points due to changes in relative demand for goods and services. This is larger than the 5 percentage point increase in the skill premium that we would attribute to changes in composition of demand using their methods on our simulated data.

Second, sectoral movements are more important in our analysis because of the way that we measure differences in skill intensity across sectors. Focusing on the EUKLEMS data, our method finds a 24 percentage point difference in factor intensities across our two sectors

[21] Bound and Johnson adjust for the increase in relative supply of high-skilled labor without accounting for the fact that the relative wage nevertheless rose. This appears to account for their much lower estimate than KM.

on average, in line with the difference in factor payments, while KM's method finds only a 4 percentage point difference. To construct efficiency unit stocks from heterogeneous labor forces, the KM method is to construct labor stocks directly using only observed labor supply measures (i.e., number of body-hours by education, experience, gender weighted by their average relative wages across industries and time). This rules out *unobserved* variation in labor quality across industries and over time. In contrast, we impute our labor supply measures indirectly from data on relative wages and industry-specific factor payment shares. Effectively, factor payment shares vary substantially more across industry than the ratios of simple bodies of different types, so that average wages of observable groups also vary substantially across industries. Our analysis assumes that all labor is paid its marginal product. We therefore attribute the discrepancy in average wages across industries for observably similar labor types to unobserved quality differences. We rule out other sources like industry rents or compensating differentials, for example. Applying the KM local approximation together with the KM measures of efficiency units on our EUKLEMS data, we account for only a 2 percentage point increase in the skill premium due to changes in the composition of demand in this period. This is slightly lower than their 5 percentage point increase because of our coarser observed grid on labor (18 types of labor rather than 320 types in KM).

In summary, there are two important factors that explain why we find a substantially larger role for compositional effects in accounting for increases in the skill premium relative to the earlier literature. The first is that our structural approach allows us to precisely disentangle the role of different driving forces. The second is that we allow for differences in unobserved productivity for workers across sectors, which in turn implies larger differences in skill intensity across sectors, thereby increasing the potential impact of compositional changes on the relative demand for skill.

2.6 Decomposing Changes in Relative Prices

While our main focus has been to understand the relative importance of different factors in generating the observed changes in the skill premium, our model also allows us to assess the importance of different factors in generating the change in the relative price of skill intensive services over time. In particular, our model suggests two distinct channels at work. As is standard in the literature on structural change with uneven technological progress across sectors, differential growth in sectoral TFP will lead to changes in relative sectoral prices. But our model also features an additional channel: because the sectors have different factor shares, changes in the relative price of factors will also lead to changes in relative sectoral prices. In particular, since the high-skill intensive sector uses skilled labor more intensively, any increase in the relative price of skilled labor will lead to a higher relative price for this sector. This effect was previously documented in equation (2.3).[22]

Here we perform some counterfactuals in our benchmark specification (i.e., $\rho = 1.42$ and $\varepsilon = 0.20$) to assess the relative importance of these two forces. In the data, the relative price of high-skill intensive services increases by 62 percentage points between 1977 and 2005, and by virtue of our calibration procedure, our model perfectly accounts for this increase. To assess the pure role played by the increase in the skill premium, we compute the implied relative price from equation (2.3) assuming that all technology parameters remain fixed at their 1977 values, but letting the skill premium increase from 1.41 to 1.90, as in the data. The result is an increase in the relative price of skill intensive services of 11 percentage points, or roughly 18% of the overall increase. In interpreting this magnitude it is important to recall our earlier discussion of the possibility that estimates of the change in relative prices are biased upward due to a failure to properly control for quality increases in the service sector. If the true change in relative prices was indeed only half as large as in the data, then the

[22] Buera and Kaboski (2012) highlight this effect in a theoretical model in which the difference in skill-intensity across the goods and service sectors arises endogenously.

change in the skill premium would account for 36% of the overall change. While still not the dominant factor, this suggests that changes in the skill premium may well be a significant factor behind changes in relative prices.

The issue of price mismeasurement notwithstanding, the direct effects of technological change are the dominant force behind the increase in the relative price of skill intensive services in our benchmark calibrated model. Moreover, it is the difference in sectoral TFP growth rates that drives this direct effect. To see this, take equation (2.3), hold the skill premium and sectoral TFPs constant and consider the pure effect of skill biased technological change. The result is that the relative price of skill intensive services would have *decreased* by 19 percentage points.

This last calculation examined the direct effect of changes in skill-biased technological change, but without incorporating the general equilibrium effect on wages. Our previous counterfactuals (see row 5 in Table 5) argued that if we eliminated changes in sectoral TFP, so that skill-biased technical change was the the only source of technological change, the skill premium would have increased from 1.41 to 1.59. If we include this effect in combination with the direct effect of skill biased technological change, the result is that the relative price of skill intensive services decreases by 15 percentage points. We conclude that skill-biased technological change is not a source of increases in the relative price of services.

In summary, we conclude that although increases in the skill premium may directly account for a non-trivial share of the increase in the relative price of high-skill intensive services, the dominant factor behind this increase is the relatively slow sectoral TFP growth in this sector.

2.7 Cross Country Analysis

In this section we extend our analysis to nine other OECD countries for which the available data exists and thereby address two distinct issues. The first issue concerns model validation, and the second issue is to assess the importance of skill-biased structural change for a larger set of countries.

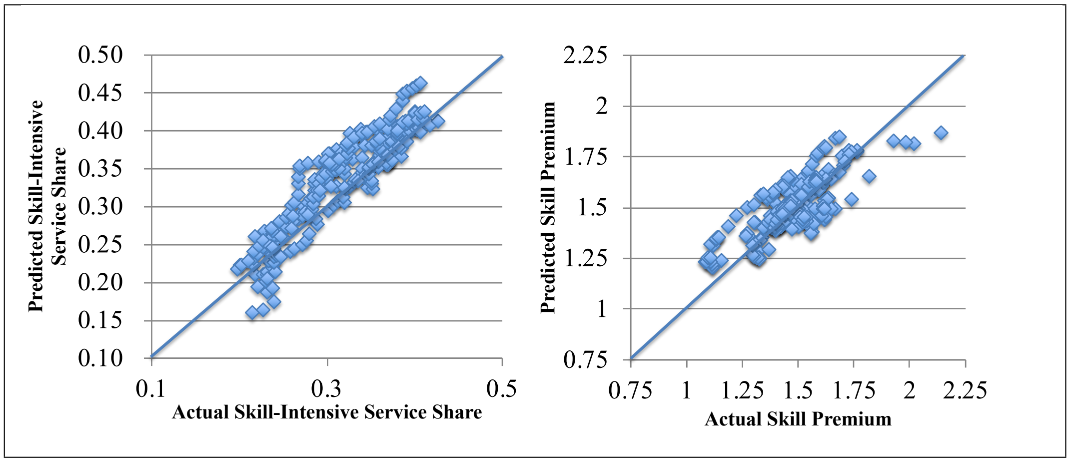
2.7.1 Model Validation Using Cross-Country Data

Our calibration procedure assigned parameter values by targeting the same number of moments as there were parameters. While both the production structure and our method for inferring technological change are very standard, we inferred values for utility function parameters by requiring that the model match the beginning and final values for sectoral valued added shares. If our utility function were mis-specified in an important way, this procedure would still allow us to fit the initial and final sectoral value added shares, but in this case we might be wary of using our calibrated specification for the counterfactual exercises.

One simple test of the specification is to consider its ability to fit not only the two endpoints of our sample, but also the entire time series. Unfortunately this is not a very stringent test for the period we are studying, since the key series in our analysis are fairly linear, and the model is able to match them fairly well.

As a somewhat more stringent test, we turn to cross country data. For this exercise we use data from the following nine countries: Australia, Austria, Denmark, Germany, Italy, Japan, the Netherlands, Spain and the United Kingdom. We assume that the utility function for each country is the same as the one implied by our benchmark calibration with $\rho = 1.42$ and $\varepsilon = 0.20$, i.e., we impose the implied values for a_G and $\bar{c}_S$. Additionally, we assume that ρ is the same for all countries. However, using the same procedures as earlier, for each country we measure the relative supply of skilled labor from the data and we use our model to infer

Figure 2.3: Model Fit in a Panel of Countries
Structural Change (left panel) and the Skill Premium (right panel)



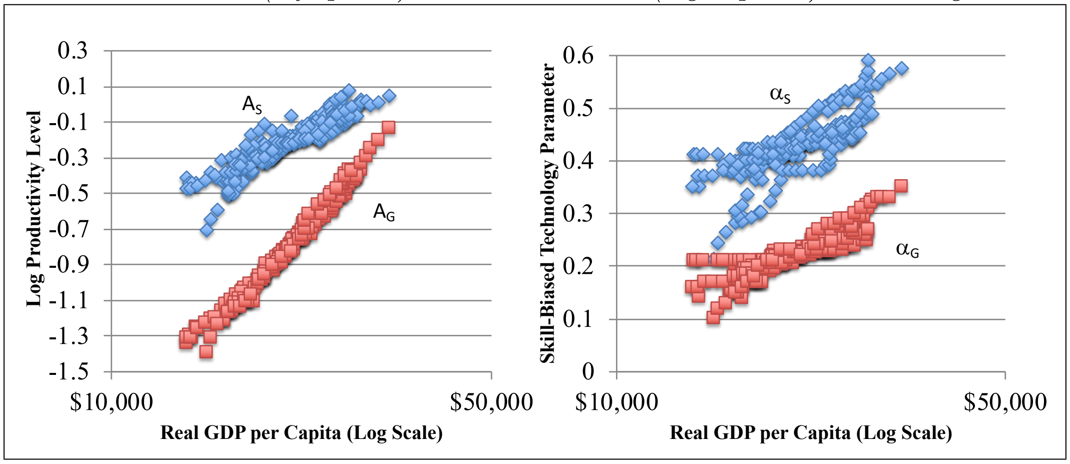
the time series for technological change. Because preference parameters are imported from the calibration using US data, we have not imposed that the model will fit the time series of interest for each country. Nonetheless, Figure 3 shows that this specification provides a reasonably good fit to the actual data for this set of countries. Because the behavior of the skilled labor share and the skill premium do differ across countries, we believe that this finding is supportive of our parsimonious structure.[23]

It is of interest to note that the above procedure implies processes for technological change that are broadly similar across countries, as shown in Figure 4.[24] To the extent that we believe the process of technology adoption and diffusion are at least generally similar across rich countries, we would view it is as somewhat problematic if our procedure indicated dramatically different processes across these countries.

[23] We note that our model does not do such a good job of matching the series for Korea. Notably, it specifically fails for the early part of the sample in which Korea has very low GDP. We intepret this as evidence that our specfiication of non-homotheticities is probably best viewed as an approximation that holds in a restricted range of incomes, and that if one wants to consider a much larger range of incomes then it is probably important to consider more general specifications such as those in Boppart (2014) and Comin et al. (2015).

[24] The plots in Figure 4 have removed country fixed effects in order to focus on the changes in technology over time rather than the cross-sectional differences.

Figure 2.4: Calibrated Technological Processes
Sector-Biased (left panel) and Skill-Biased (right panel) Technologies



Note: The diamonds (squares) correspond to the high (low) -skill intensive sector

2.7.2 Skill-Biased Structural Change and the Skill Premium in Cross Country Data

In this subsection we assess the extent to which skill biased structural transformation has influenced the skill premium in each of the countries studied in the previous subsection. We could carry out this calculation for the specifications in the last subsection, i.e., assuming the same preference parameters for these countries as in the US. A potential disadvantage of this method is that although the model with common preference parameters across countries offers a good fit to the cross country time series data, it does not necessarily account for all of the changes in the skill premium for each of the countries. Alternatively, we could assume country specific values for a_G and $\bar{c}_S$ and simply repeat the analysis that we have carried out for the US for each of the additional economies. These two methods provide fairly similar answers, and in the interest of space we only report the results of the second exercise, which are shown in Table 8.

To compute the values in Table 8 we first calculate the contribution of all forms of technological change by computing the difference between the actual skill premium in 2005 versus the skill premium that would have existed in 2005 if there had been no technological change relative to 1977 but allowing for the observed change in the supply of skill. We then isolate the fraction of this overall contribution of technological change that is due to skill biased structural change by computing the fraction of this change that is accounted for by changes in the A_j's. Relative to our earlier calculations, this procedure corresponds to the change moving from row 3 to row 4 in Table 5.

Table 2.8: Contribution of SBTC Across Countries

Australia	0.15
Austria	0.28
Denmark	0.14
Spain	0.25
Germany	0.24
Italy	0.47
Japan	0.17
Netherlands	0.21
UK	0.32

The magnitude of this share varies significantly, from a low of 14% in Denmark to a high of 47% in Italy, but the mean value of 25% is very much in line with our estimates from the US. We conclude that the demand side forces associated with skill biased structural change seem to be quantitatively significant in a broad group of advanced economies.

2.8 Conclusion

Using a broad panel of advanced economies, we have documented a systematic tendency for development to be associated with a shift in value added to high-skill intensive sectors. It follows that development is associated with an increase in the relative demand for high skill workers. We coined the term skill-biased structural change to describe this process. We have built a simple two-sector model of structural transformation and calibrated it to US data over the period 1977 to 2005 in order to assess the quantitative importance of this mechanism for understanding the large increase in the skill premium during this period. We find that technological change overall increased the skill premium by roughly 100 percentage points, and that between 20 and 27 percent of this change is due to technological change which operated through compositional changes.

Our findings have important implications for predicting the future evolution of the skill premium, since the continued growth of the value added share of the high-skill intensive sector will exert upward pressure on this premium even in the absence of skill-biased technological change.

In order to best articulate the mechanism of skill-biased structural change we have purposefully focused on a simple two-sector model. As we noted in Section 2, there is good reason to think that the mechanism we have highlighted is also at work at a more disaggregated level, so it is of interest to explore this mechanism in a richer model. The early literature has also emphasized the possibility that increases in trade might lead to changes in the composition of valued added across sectors. Katz and Murphy (1992b) specifically noted this possibility, and more recent analyses include Feenstra and Hanson (1999) and Autor et al. (2013). We think it is important to note that the compositional effects we have focused on are not likely to be reflecting changes due to trade. The reason for this is that our high-skill intensive sector is composed entirely of industries from the service sector. It is plausible that part of what we identified as within sector skill biased technical change may at least in part reflect compositional effects due to trade, to the extent that trade had caused manufacturing activity in the US to shift to more skill intensive industries.

Chapter 3

Natural Disasters and Growth: The Role of Foreign Aid and Disaster Insurance

Rodolfo Manuelli Juan Ignacio Vizcaino [1]

3.1 Introduction

The immediate economic impact of a natural disaster strike is large and negative. The empirical literature has found mixed results about the longer run effect of natural disasters on economic activity. To the extent that global climate change will likely increase the prevalence of some forms of natural disasters it is important to develop a framework that is suitable to interpret the evidence at the same time that provides some guidance on the effect of policies.

What is the evidence? In the last few years there has been extensive empirical research on the economic impact of natural disasters on growth. A cursory reading of the literature

[1] Manuelli: Department of Economics, Washington University in St. Louis, One Brookings Drive, St. Louis, MO 63130 (email: *manuelli@wustl.edu*); Vizcaino: Department of Economics, Washington University in St. Louis, One Brookings Drive, St. Louis, MO 63130 (email: *jivizcaino@wustl.edu*)

suggests that there is a significant disagreement about the short and long run consequences of natural disaster strikes. In some cases the evidence points to a positive relationship between the risk of natural disasters and economic growth. For example, Skidmore and Toya (2007) find that the average number of natural disasters is positively correlated with growth. Kousky (2014) reviews a large number of studies and finds that natural disasters have a modest impact on economic activity. At the other end, Hsiang and Jina (2014), Berlemann and Wenzel (2016) and Bakkensen and Barrage (2017) reach the opposite conclusion: a natural disaster strike -in these papers the analysis is restricted to tropical cyclones- decreases the growth rate, and the impact is relatively long lasting. We view these differences as indicating not only that the measurement of a natural disaster event is difficult and mired with error but also that it is necessary to take into account heterogeneity across countries in the activities that can influence the effect of a natural disaster.

There are two other pieces of evidence that seem relevant to motivate what a theoretical model should include. First, Berlemann and Wenzel find that the growth impact of a natural disaster varies depending on the country's level of development: a natural disaster strike in a relatively rich country has very small growth effects while a similar event in a poor country results in large decreases in growth .[2] Second, von Peter et al. (2012) find that the growth impact of a natural disaster strike depends on whether the loss was insured or not: insured losses do not appear to have a significant impact on growth while uninsured losses have a negative impact.[3] It is not clear that these two are independent observations as it is possible that high income countries are also countries that are better insured against natural disasters. Studying the impact of hurricanes in the U.S. Deryugina (2017) finds that following a strike the affected area receives transfers -emergency aid and insurance payments- in an amount

[2] Hsiang and Jina on the other hand find no significant differences between countries with above the median income and countries below. This way of categorizing rich and poor is possibly too coarse to get significant results.

[3] The result must be interpreted with care since measured losses are, at best, a very noisy proxy for actual losses.

close to the estimates of the damages caused by a hurricane. She finds that in her sample hurricanes have a negligible impact on income. That is, transfers that compensate for the loss result in no growth effects. We read this research as suggesting that the role of transfers and insurance should not be neglected when analyzing the economic impact of natural disasters.

In this paper we develop a continuous time stochastic growth model that is rich enough to account for the evidence. We study optimal consumption and investing under alternative market structures. We explore the role of foreign disaster aid and study its impact on saving decisions as well as the choice of sectoral allocation of investment. We also explore the impact upon endogenous decisions of the availability of actuarially fairly priced disaster insurance. Finally, we explore the effect, both in terms of growth and welfare, of delays in the provision of aid and in the payouts of insurance contracts.

Not surprisingly given the existing results on stochastic growth models, the growth impact of shocks depends on the curvature of the utility function, even though the Poisson shocks that we use to capture the large and unusual natural disaster shocks are not of the more standard variety.

Some of the less intuitive results include:

- Foreign aid received when the natural disaster impacts a country in the normal phase and aid when the country is in the disaster regime (roughly experiences two events within a short time) have potentially opposite impacts on the growth rate during the recovery period.

- Under reasonable conditions on the prevention and mitigation technologies, increases in foreign aid reduce investment in mitigation activities and, as a result, delay the recovery from the disaster (i.e. increase the expected duration of the low productivity regime).

- Depending on parameters, foreign aid can either crowd out the demand for insurance or induce a country to "over insure." It is even theoretically possible that a country becomes a net seller of catastrophe insurance. This can happen if the country expects a large inflow of foreign aid contingent on a natural disaster strike since it can use the reverse insurance to increase consumption in normal times. Even though we do not expect this to be the outcome under a realistic calibration, the possibility shows the incentives that must be taken into account when creating a market for disaster insurance.

- Increased frequency of natural disasters has a growth effect even holding the expected loss of stocks from a strike constant. This simply illustrates the non-linearity and extensive cross equations restrictions implies by the theory. Moreover it shows that measurement matters since expected losses and frequency can have opposite growth effects.

In order to make progress quantifying the impact of natural disasters we conduct a quantitative exercise. We pick parameters to match the evidence on the effect of cyclones and we asses the effect of foreign aid, insurance and improvements in prevention and mitigation technologies (in progress).

The paper closest to ours is Bakkensen and Barrage. They also analyze a growth model under normality assumptions and note the difference between natural disaster risk and strikes. The main difference is that we emphasize the role of foreign aid, insurance markets and prevention and mitigation technologies. In addition, our model allows for the possibility of higher and or lower growth rate in the post impact period while theirs implies that the growth rate is unchanged.

Our work is also related to the literature on the macro impact of large shocks which includes Barro (2009), Jones and Olken (2008), Gourio (2012) and Gabaix (2012).

In section 3.2 we describe the basic model and study separately the equilibrium allocations in the case in which the country does not have access to catastrophe insurance and the case in which it does. Section 3.3 contains the quantitative analysis and section 3.4 offers some conclusions.

3.2 Model

We study an economy populated by a representative dynasty. Since we abstract away from externalities the competitive equilibrium coincides with the solution to the planner's problem. We assume that the economy is closed except for limited access to a disaster insurance market. The model includes two types of shocks: standard TFP shocks and natural disaster shocks that are modeled as Poisson arrivals.

We view the economy as being in one of two regimes. The *normal* regime is the high productivity regime while the *disaster* regime is associated with low productivity. An economy that is in the normal regime switches to the disaster regime upon receiving a natural disaster strike. It reverts back to the normal regime with an instantaneous probability that depends on resources allocated to recover.

On the technology side we consider a standard two capital good Ak model with the following capital accumulation technologies:

$$dk = Akdt + \sigma kdZ_t - \left(1 - \mu_k^\delta(\kappa)\delta_k\right) kdN_t,$$

$$dh = Hhdt - \left(1 - \mu_h^\delta(\kappa)\delta_h\right) hdN_t,$$

where Z_t is a standard Brownian motion and N_t is a Poisson with parameter η. A realization of this Poisson corresponds to a natural disaster strike. The term $\mu_j^\delta(\kappa)\delta_j$ measures the amount of j-type of capital that is available after the natural disaster. We assume that

$\mu_j^\delta(\kappa)\delta_j \in (0,1)$. In this context $\mu_j^\delta(\kappa)$ is the average recovery rate and δ_j is random and has mean one. This component is meant to capture uncertainty about the strength of the natural disaster.

As it is standard in this Merton-type models we assume that $A > H$ but that the parameters are such that the share of both capital stocks in total wealth is strictly between zero and one.

We assume that the country can spend resources in activities that reduce the impact of a hurricane. For example, sea walls and better construction standards can significantly reduce the effect of a cyclone in a coastal area. In addition, we allow for the possibility that the country purchase insurance. In the simplest version of the model we introduce a standard insurance contract purchases from the rest of the world: the country pays a premium conditional on no natural disasters occurring and receives a payment when there is a hurricane strike.

In the simple version of the model we let total wealth be denoted by w, with $w = k + h$. Then the law of motion of wealth in the normal phase is

$$
\begin{gathered}
dw = \left[\left(\alpha A + (1-\alpha) H\right) - (\kappa + b + c)\right] wdt \\
+\sigma \alpha w dZ_t - \left(1 - \mu_k^\delta(\kappa)\delta_k\right) \alpha w dN_t - \left(1 - \mu_h^\delta(\kappa)\delta_h\right)(1-\alpha) w dN_t.
\end{gathered}
$$

In this specification $k = \alpha w$ and $h = (1-\alpha)w$, κw is the total amount of resources allocated to prevention, bw is the premium corresponding to the insurance contract and cw is aggregate consumption. We assume that $\mu_j^\delta(\kappa)$ is increasing in κ.

The occurrence of a natural disaster has several effects. First, it causes a jump in the level of wealth. Let w' be the stock of wealth after the strike. Then,

$$w' = w \left[\underbrace{\mu_k^\delta(\kappa)\delta_k(1+\zeta_k)\alpha + \left(\mu_h^\delta(\kappa)\delta_h(1+\zeta_h)\right)(1-\alpha) + I(b) + \zeta_w}_{\equiv \beta(\alpha,b,\kappa,\mu^\delta,\delta,\zeta)} \right]$$

The term $\beta(\alpha, b, \mu^\delta, \delta, \zeta)$ captures the loss of wealth associated with the arrival of a natural disaster. For each capital of type j the post-hurricane level is $\mu_j^\delta(\kappa)\delta_j(1+\zeta_j)$ of the pre-hurricane level. The term ζ_j captures the amount of capital specific foreign aid post-hurricane.[4] The term $I(b)$ is the payoff per unit of wealth of the insurance contract. Finally ζ_w stands for foreign aid that can be used at the discretion of the country. In general, we expect that the sum of all the terms except for the insurance payout will be less than one.

Second, we assume the occurrence of a natural disaster event is associated with lower the productivity of both types of capital but the loss is a function of both the resources allocated to prevention, as measured by κ as well as resources destined to mitigation which we denote by κ_D.[5] Thus, as a matter of notation we assume that post-strike the productivities are

$$\begin{aligned} A_D &= A(\kappa, \kappa_D) \leq A, \\ H_D &= H(\kappa, \kappa_D) \leq H. \end{aligned}$$

In what follows, to ease notation, we will be using (A_D, H_D) without explicitly noting their dependence on (κ, κ_D).

We assume that the duration of this low productivity phase is endogenous and depends on the amount of "mitigation resources" spent by the country. We assume that the switch back to normal times is well described by a Poisson process M_t with parameter $\upsilon(\kappa_D)$. Since the expected duration of the low productivity phase following a natural disaster strike is

[4] In a future extension we will also consider the possibility that the insurance payout is capital specific.

[5] We use the subscript D to denote the relevant values in the disaster regime.

$1/\upsilon(\kappa_D)$ we assume that $\upsilon(\kappa_D)$ is increasing in κ_D.

The corresponding feasibility constraint during the disaster phase is given by

$$dw = \left[\left(\alpha_D A_D + (1-\alpha_D) H_D\right) - \left(\kappa_D + c_D + b_D\right)\right] wdt$$
$$+\sigma\alpha_D w dZ_t + wdM_t - \left(1-\mu_k^\delta(\kappa)\delta_k\right)\alpha_D w dN_t - \left(1-\mu_h^\delta(\kappa)\delta_h\right)(1-\alpha_D) w dN_t,$$

where the last two terms capture the loss associated with another natural disaster strike while the country is still in the disaster phase.[6]

We assume that the utility function is given by

$$u(C) = c^{1-\gamma}\frac{w^{1-\gamma}}{1-\gamma}, \text{ where } C = cw.$$

Let the value function in normal (disaster) times be $V_N(w)$ ($V_D(w)$). Given the linearity in the technology and the assumption that preferences are of the CRRA variety we conjecture that

$$\begin{aligned} V_N(w) &= V_N\frac{w^{1-\gamma}}{1-\gamma}, \\ V_D(w) &= V_D\frac{w^{1-\gamma}}{1-\gamma}. \end{aligned}$$

The HJB equations of the planner's problem (which coincides with the competitive allocation) are,

$$\rho V_N\frac{w^{1-\gamma}}{1-\gamma} = \max_{c,\alpha,\kappa,b}\left\{c^{1-\gamma}\frac{w^{1-\gamma}}{1-\gamma} + V_N w^{1-\gamma}\left[\left(\alpha A + (1-\alpha) H\right) - (\kappa + b + c)\right]\right. \tag{3.1}$$
$$\left. -\gamma V_N w^{1-\gamma}\frac{\sigma^2}{2}\alpha^2 + \eta\left[V_D\frac{w^{1-\gamma}}{1-\gamma}E\left[\left(\beta(\alpha,b,\kappa,\mu^\delta,\delta,\zeta)\right)^{1-\gamma}\right] - V_N\frac{w^{1-\gamma}}{1-\gamma}\right]\right\}.$$

[6]To keep the model stationary we assume that if there is another strike while the economy is in the disaster phase there is no further decrease in productivity. The only impact of this "second" strike is to reduce the stocks of both types of capital.

The first three terms on the right hand side are standard. The last term captures the value loss associated with a natural disaster strike.

The corresponding equation for the disaster case is

$$\rho V_D \frac{w^{1-\gamma}}{1-\gamma} = \max_{c,\alpha,\kappa_D,b_D} \left\{ c^{1-\gamma}\frac{w^{1-\gamma}}{1-\gamma} + V_D w^{1-\gamma}\left[\left(\alpha A_D + (1-\alpha) H_D\right) - (\kappa_D + c + b_D)\right] \right. \tag{3.2}$$
$$-\gamma V_D w^{1-\gamma}\frac{\sigma^2}{2}\alpha^2 + \upsilon(\kappa_D)\left[V_N \frac{w^{1-\gamma}}{1-\gamma} - V_D \frac{w^{1-\gamma}}{1-\gamma}\right]$$
$$\left. +\eta\left[V_D \frac{w^{1-\gamma}}{1-\gamma} E\left[\left(\beta(\alpha, b_D, \kappa, \mu^\delta, \delta, \zeta_D)\right)^{1-\gamma}\right] - V_D \frac{w^{1-\gamma}}{1-\gamma}\right]\right\}.$$

In addition to the standard terms corresponding to TFP shocks the value of the problem depends on the gain associated with switching to the normal regime (and this is driven by a Poisson with parameter $\upsilon(\kappa_D)$) as well as the potential loss associated with another natural disaster hitting the economy while it is still in the disaster phase (and this is driven by an independent Poisson with parameter η). In this formulation we allow the amount of aid conditional on a natural disaster strike and the country being in the disaster phase, ζ_D, to be potentially different from ζ.[7]

3.2.1 No Disaster Insurance

It seems useful to understand the forces at work to consider a sequence of versions that simplify the problem by shutting down several channels. Let us first consider that case in which the loss of stocks associated with natural disaster is small. To capture this we set $\delta_k = \delta_h = 1$. We also take for now the choice of investment in prevention and mitigation as exogenous and assume no insurance.

Given that the utility function is unbounded it is clear that existence depends on pa-

[7]Since on average a country that experiences another event while still in the disaster regime corresponds to a country that has been hit twice in a relatively short time by a natural disaster we allow for donors to respond differentially.

rameter values. Thus, until we get to the quantitative section of the paper we will simply assume existence of an equilibrium. Put it differently the model only makes economic sense for those parameter values consistent with existence of an equilibrium.

To economize on notation we define

$$\begin{aligned}
P &= H + \alpha_N(A-H) - \gamma\frac{\sigma^2}{2}\alpha_N^2 - \kappa, \\
P_D &= H_D + \alpha_D(A_D-H_D) - \gamma\frac{\sigma^2}{2}\alpha_D^2 - \kappa_D. \\
V &= \frac{V_N}{V_D}
\end{aligned}$$

It is understood that P depends on κ and the other variables that affect α_N and the same applies to P_D even though we do not make that dependence explicit.

Proposition 1. *Let* (α_N, V) *be the unique solution to the following equations*

$$\begin{aligned}
\alpha_N &= \frac{A-H}{\gamma\sigma^2} + \frac{\eta\Delta_k(\kappa)}{V\beta(\alpha_N, b, \kappa, \mu^\delta, \delta, \zeta)^\gamma}\frac{1}{\gamma\sigma^2}, \\
\alpha_D &= \frac{A_D-H_D}{\gamma\sigma^2} + \frac{\eta\Delta_k(\kappa)}{\beta(\alpha_D, b_D, \kappa, \mu^\delta, \delta, \zeta_D)^\gamma}\frac{1}{\gamma\sigma^2}
\end{aligned} \tag{3.3}$$

where

$$\Delta_k(\kappa) = \mu_k^\delta(\kappa)(1+\zeta_k) - \mu_h^\delta(\kappa)(1+\zeta_h)$$

and

$$\begin{aligned}
&\rho + \eta + \upsilon(\kappa_D) - (1-\gamma)P_D - \eta\beta(\alpha_D, b_D, \kappa, \mu^\delta, \delta, \zeta_D)^{1-\gamma} - \upsilon(\kappa_D)V \\
= \; &\left[(\rho+\eta) - (1-\gamma)P\right]V^{1/\gamma} - \eta\beta(\alpha_N, b, \kappa, \mu^\delta, \delta, \zeta)^{1-\gamma}V^{\frac{1-\gamma}{\gamma}}.
\end{aligned} \tag{3.4}$$

The expected growth rates in each regime conditional on no regime change are

$$\mu^N = H + \alpha_N(A-H) - \kappa - V_N^{-1/\gamma} \tag{3.5}$$

and

$$\mu^D = H_D + \alpha_D (A_D - H_D) - \kappa_D - V_D^{-1/\gamma}, \tag{3.6}$$

where

$$\gamma V_D^{-1/\gamma} = \rho + \eta + \upsilon(\kappa_D) - (1-\gamma)P_D - \eta\beta(\alpha_D, b_D, \kappa, \mu^\delta, \delta, \zeta_D)^{1-\gamma} - \upsilon(\kappa_D)V, \tag{3.7}$$

and

$$\gamma V_N^{-1/\gamma} = \rho + \eta - (1-\gamma)P - \eta\frac{\beta(\alpha_N, b, \kappa, \mu^\delta, \delta, \zeta)^{1-\gamma}}{V} \tag{3.8}$$

Proof. (See Appendix) □

In a standard Merton portfolio problem the share of risky assets in the portfolio is given by

$$\frac{A-H}{\gamma\sigma^2},$$

since both α_j equal the share as prescribed by the Merton result plus a term whose sign depends on the sign of $\Delta_k(\kappa)$ it follows that when the expected capital loss associated with the k-type of capital exceeds that of the h-type of capital then the optimal α_j falls short of the Merton value and the opposite is true when the values are reversed.

This result highlights one of the channels that, in the model, can account for the difference between μ^N and μ^D. First, the fact that $A_D < A$, and $H_D < H$ implies that $\mu^D < \mu^N$. However, there are two other forces that can, potentially, reverse this. First, there is the standard saving effect captured by $V_N^{-1/\gamma}$ and $V_D^{-1/\gamma}$. In this case the reason why saving might be lower in the normal regime is that, starting from that phase, the economy will have lower returns if it switches to a disaster phase, while this is not the case if it is already in the disaster regime. Of course for this effect to dominate the income effect it is necessary that the utility function display relatively low curvature. Second, it is possible that $\alpha_D > \alpha_N$ and this increase in the share of the high return capital can increase growth. Whether this happens or not depends in complicated ways on the details of the distribution of foreign aid (ζ, ζ_D) among other effects.

3.2.2 Disaster Insurance and the Role of Aid

In this section we assume that the country has access to an insurance market and that insurance is fairly priced by the rest of the world. We do not need to assume that the interest rate is the same as the domestic discount rate. In this case it follows that zero profits in this activity implies that the relationship between premiums and payoffs are

$$I(b) = \frac{b}{\eta}.$$

In this case, it is possible to show that the optimal choice of insurance during the normal phase is such that[8]

$$V\beta(\alpha_N, b, \kappa, \mu^\delta, \delta, \zeta)^\gamma = 1,$$

and the share of the portfolio allocated to the risky asset is

$$\alpha_N = \frac{A - H}{\gamma\sigma^2} + \eta\Delta_k(\kappa).$$

During the disaster phase the optimal choice is

$$\beta(\alpha_D, b_D, \kappa, \mu^\delta, \delta, \zeta_D) = 1$$

which corresponds to full insurance.

[8]This condition does not imply that the post-strike level of wealth is lower than the pre-strike. In particular, one can show that if $\gamma > 1$ then $V < 1$ and hence $\beta(\alpha_N, b, \kappa, \mu^\delta, \delta, \zeta) > 1$ which implies more post transfer wealth.

Since the productivity of aggregate capital (or wealth) is given by

$$H + \alpha_N(A - H) = H + \frac{(A - H)^2}{\gamma\sigma^2} + \eta\Delta_k(\kappa)\,(A - H)$$

then the type of disasters that result in larger losses for physical than human capital — corresponding to $\Delta_k(\kappa) < 0$— increases in the frequency of natural disasters (i.e. an increase in η) decrease the aggregate productivity as it results in a smaller investment in the high return (and high loss in the event of a natural disaster) capital. Of course, if $\Delta_k(\kappa) > 0$ the same forces result in higher productivity.

In order to study the effects of insurance it is convenient to emphasize the version of the model in which the properties of the natural disaster do not directly affect the composition of the portfolio. To be precise, we assume that $\Delta_k(\kappa) = 0$. In this case,

$$\begin{aligned} \alpha_N &= \frac{A - H}{\gamma\sigma^2}, \\ \alpha_D &= \frac{A_D - H_D}{\gamma\sigma^2} \end{aligned}$$

are independent of natural disasters and

$$\beta(\alpha, b, \kappa, \mu^\delta, \delta, \zeta) = \mu^\delta(\kappa) + \frac{b}{\eta} + \zeta,$$

where

$$\mu^\delta(\kappa) = \mu_k^\delta(\kappa)(1 + \zeta_k) = \mu_h^\delta(\kappa)(1 + \zeta_h).$$

In our notation we distinguish between the effect of foreign aid when the natural disaster strike occurs during a normal phase, which we denoted by ζ, from the case in which the transfer follows a natural disaster strike that occurs when the country is already in the disaster phase, which we denoted by ζ_D.

In some cases, it is useful to consider the case $\zeta = \zeta_D = \bar{\zeta}$ which assumes that foreign aid is not contingent on whether the country had been recently affected by another natural disaster.

It follows that

$$P = H + \frac{(A-H)^2}{2\gamma\sigma^2} - \kappa, \tag{3.9}$$

and

$$P_D = H_D + \frac{(A_D - H_D)^2}{2\gamma\sigma^2} - \kappa_D \tag{3.10}$$

are independent of properties of natural disasters and, hence, can be taken as given

The following proposition summarizes the basic implications of the model for $V = V_N/V_D$, the relative value of the problem in the normal and disaster phases in the absence of insurance.

Proposition 2 (Relative Valuation Under No Insurance)**.** *The relative value of the problems $V = V_N/V_D$ solves the following equation:*

$$\begin{aligned} & (\rho+\eta) - (1-\gamma)P - \eta\frac{\beta(\alpha, b, \kappa, \mu^\delta, \delta, \zeta)^{1-\gamma}}{V} \\ = \; & \left[\rho + \eta + \upsilon(\kappa_D) - (1-\gamma)P_D - \eta\beta(\alpha, b, \kappa, \mu^\delta, \delta, \zeta_D)^{1-\gamma} - \upsilon(\kappa_D)V\right] V^{-1/\gamma}. \end{aligned} \tag{3.11}$$

The solution V^* has the following properties

1. *If* $\gamma \in (0,1)$

$$\begin{aligned} \frac{\partial V^*}{\partial \zeta} &> 0, \frac{\partial V^*}{\partial \zeta_D} < 0 \\ \frac{\partial V^*}{\partial \upsilon(\kappa_D)} &< 0, \frac{\partial V^*}{\partial \eta} \Big|_{\zeta=\zeta_D} < 0 \end{aligned}$$

2. if $\gamma > 1$

$$\begin{aligned} \frac{\partial V^*}{\partial \zeta} &< 0, \frac{\partial V^*}{\partial \zeta_D} > 0, \\ \frac{\partial V^*}{\partial \upsilon(\kappa_D)} &> 0, \frac{\partial V^*}{\partial \eta} \mid_{\zeta=\zeta_D} > 0 \end{aligned}$$

Proof. (see the Appendix) □

The average growth rates in each regime, contingent on no strikes, are given by

$$\begin{aligned} \mu^N &= \frac{1}{\gamma}\left[P + \eta\left(\mu^\delta(\kappa) + \zeta\right) - (\rho + \eta)\right] \\ \mu^D &= \frac{1}{\gamma}\left[\left(P_D + \eta\left(\mu^\delta(\kappa) + \zeta_D\right)\right) - (\rho + \eta + \upsilon(\kappa_D)) + \upsilon(\kappa_D)V\right]. \end{aligned}$$

We can now summarize the impact of foreign aid and properties on natural disaster on the growth rate in each phase. Not surprisingly, the qualitative implications depend on whether the country has access to insurance

Proposition 3 (Full Insurance: The Effect of Foreign Aid)**.** *The impact of aid when the country has access to full insurance is given by*

$$\frac{\partial \mu^N}{\partial \zeta} > 0, \frac{\partial \mu^N}{\partial \zeta_D} = 0, \frac{\partial \mu^N}{\partial \bar{\zeta}} > 0.$$

and

$$\frac{\partial \mu^D}{\partial \zeta} < 0, \frac{\partial \mu^D}{\partial \zeta_D} = 0, \frac{\partial \mu^D}{\partial \bar{\zeta}} < 0.$$

Proof. (see the Appendix) □

Proposition 4 (No Insurance: The Effect of Foreign Aid)**.** *The impact of aid when the* *country has access to full insurance is given by*

$$\frac{\partial \mu^N}{\partial \zeta} < 0, \frac{\partial \mu^N}{\partial \zeta_D} > 0.$$

and

$$\frac{\partial \mu^D}{\partial \zeta} > 0, \frac{\partial \mu^D}{\partial \zeta_D} = \textit{indeterminate}.$$

Proof. (see the Appendix) □

The model implies potentially heterogenous growth effects of foreign aid depending on the specific details of how it is awarded. Increases in regime neutral aid, as measured by $\bar{\zeta}$, unambiguously increase the growth rate in the normal phase at the same time that is decreases the growth rate in the disaster phase.

Increases in (promised) aid when the country experiences a natural disaster strike but is otherwise in a normal phase, that is, increases in ζ, have opposite effects on the growth rate on the two phases.

Next we study how the different dimensions of a natural disaster affect growth. Given the relatively simple model that we study we can summarize the relevant dimensions as

- Frequency of strikes: $1/\eta$.
- Duration of the disaster phase: $1/\upsilon(\kappa_D)$
- Loss of stocks: $\mu^\delta(\kappa)$
- Loss of productivity: $P_D/P = 1 - \phi$

Proposition 5 (Growth and the Structure of Natural Disasters). *1. Changes in η*

$$\frac{\partial \mu^N}{\partial \eta} < 0, \frac{\partial \mu^D}{\partial \eta} \mid_{\gamma \in (0,1)} < 0, \frac{\partial \mu^D}{\partial \eta} \mid_{\gamma > 1} = \textit{ ambiguous}.$$

2. Changes in $\upsilon(\kappa_D)$

$$\frac{\partial \mu^N}{\partial \upsilon(\kappa_D)} = 0, \frac{\partial \mu^D}{\partial \upsilon(\kappa_D)} \mid_{\gamma \in (0,1)} > 0, \frac{\partial \mu^D}{\partial \upsilon(\kappa_D)} \mid_{\gamma > 1} < 0.$$

3. Changes in $\mu^\delta(\kappa)$

$$\frac{\partial \mu^N}{\partial \mu^\delta(\kappa)} > 0, \frac{\partial \mu^D}{\partial \mu^\delta(\kappa)} > 0.$$

4. Changes in ϕ

$$\frac{\partial \mu^N}{\partial \phi} < 0, \frac{\partial \mu^D}{\partial \phi} < 0.$$

Proof. (see the Appendix) □

Some of the results are as expected: natural disasters that result in more destruction of stocks and that are associated with lower productivity unambiguously decrease growth. However the impact of duration (or frequency) of the phenomena have less intuitive effects. Consider, for example the impact of a decrease in the duration of the low productivity phase, that is, a faster recovery. This improvement has no impact on the growth rate in the normal phase and can actually decrease the growth rate in the disaster phase. This can happen when the utility function has more curvature than the log. In this case income effects dominate and the expectation of a faster recovery (and the associated higher return to investment) does not result in higher savings. Rather, the optimal policy increases consumption.

The model is highly non-linear and it suggests that different elements of a natural disaster can have different impacts on growth. To illustrate this consider the impact of increasing the frequency of natural disasters, η, at the same time that the expected loss associated with a natural disaster is held constant. To be precise let the instantaneous return on capital be denoted z, then the total return taking into account that a fraction $1 - \mu^\delta(\kappa)$ is lost in the case of a natural disaster is simply

$$\frac{z}{\rho + \eta \left(1 - \mu^{\delta}(\kappa)\right)}.$$

Thus a measure of a natural disaster economic cost is $\eta \left(1 - \mu^{\delta}(\kappa)\right)$. We want to compare the growth impact of different natural disasters that are associated with exactly the same expected loss. Let $\eta \left(1 - \mu^{\delta}(\kappa)\right) = m$. Thus, holding m constant we want to determine the impact of more frequent (higher η) natural disasters. Thus, this captures the tradeoff between more frequent, but less destructive, events and more rare but more damaging natural disasters.

Simple algebra shows that

$$\frac{\partial \mu^{N}}{\partial \eta} \mid_{m=\bar{m}} = \zeta > 0.$$

Thus, when it comes to evaluating the growth impacts of natural disasters in normal times more frequent and less severe events are growth enhancing. The result shows that empirical work that tries to ascertain the growth effects of a natural disaster and uses expected losses as its measure of impact will get biased estimates depending on the distribution of frequencies.

The effect of the expected time in between strikes $(1/\eta)$ on the growth rate in the post-strike phase is ambiguous. Formally, the impact on μ^{D} is

$$\frac{\partial \mu^{D}}{\partial \eta} \mid_{m=\bar{m}} = \zeta_{D} + \upsilon(\kappa_{D}) \frac{\partial V}{\partial \eta} \mid_{m=\bar{m}} .$$

The sign of the term $(\partial V/\partial \eta) \mid_{m=\bar{m}}$ depends on the elasticity of substitution. When income effects dominate $(\gamma > 1)$, it is positive and more frequent natural disasters are growth enhancing. If $\gamma < 1$ the last term is negative and the whole expression would be negative if ζ_{D} is small.

The Demand for Disaster Insurance What is the optimal choice of insurance? One can show that the demand for insurance in the disaster phase is such that it completely offsets the capital loss, that is

$$b_D = \eta(1 - (\mu^\delta(\kappa) + \zeta_D)).$$

In this situation there is compete crowding out of private insurance by foreign aid. This suggests that efforts to create a market for catastrophe bonds have to take into account the negative incentives associated with the expectation of foreign aid. For sufficienlty high levels of foreign aid, high ζ_D the country will be a net seller of catastrophe bonds.

In the normal phase the demand for insurance is given by

$$b_N = \eta(V^{-1/\gamma} - (\mu^\delta(\kappa) + \zeta_D)).$$

Since the model imposes no restrictions on this demand it is possible for the country to "overinsure" in the sense that, in equilibrium, $\beta(\alpha_N, b, \kappa, \mu^\delta, \delta, \zeta) > 1$, the post-strike relative wealth can be greater than one. In fact, this is the case if $\gamma > 1$. As Proposition 2 shows in this case $V < 1$ and since the optimal choice of insurance requires that

$$V\beta(\alpha_N, b, \kappa, \mu^\delta, \delta, \zeta)^\gamma = 1$$

it follows that $\beta(\alpha_N, b, \kappa, \mu^\delta, \delta, \zeta) > 1$. The reason for this is that the country is using the insurance market to insure as well against the low productivity during the disaster phase. One way of doing that is by acquiring more wealth conditional on the shock and this is exactly the type of contract that the insurance scheme offers.

In the case that $\gamma \in (0, 1)$, then the optimal choice is such that there is incomplete insurance, that is, $\beta(\alpha_N, b, \kappa, \mu^\delta, \delta, \zeta) < 1$. Finally it is possible for the country to "sell" insurance (issuer of catastrophe bonds). This corresponds to the case

$$b_N = \eta(V^{-1/\gamma} - (\mu^\delta(\kappa) + \zeta_D)) < 0$$

which can happen when V is sufficiently large. Even though this might seem paradoxical, the key driver of this role reversal is foreign aid. If the country expects a large ζ then it chooses to increase current consumption in exchange for lower future consumption. Effectively, the country is selling some of its right to the foreign aid it will receive in the case of a natural disaster strike.

Optimal Choice of Prevention and Mitigation

In general it is not possible theoretically to determine how changes in foreign aid will affect endogenous prevention and mitigation efforts. I this section we make some progress and report some partial results. We take the objective function to maximize V_N. Thus function V_N satisfies

$$\gamma V_N^{-1/\gamma} = \rho + \eta + (\gamma - 1)\left[P + \eta\left(\mu^\delta(\kappa) + \zeta\right)\right] - \gamma\eta V^{-1/\gamma}. \tag{3.12}$$

For an interior maximum we require that $\partial V_N/\partial\kappa$ and $\partial V_N/\partial\kappa_D$ be equal to zero. Simple algebra implies that

$$\frac{\partial V_N}{\partial \kappa} = 0 \Leftrightarrow \eta V^{-\left(\frac{1}{\gamma}+1\right)}\frac{\partial V}{\partial \kappa} = (1-\gamma)\left[-1 + \eta\frac{d\mu^\delta}{d\kappa}(\kappa)\right], \tag{3.13}$$

and

$$\frac{\partial V_N}{\partial \kappa_D} = 0 \Leftrightarrow \frac{\partial V}{\partial \kappa_D} = 0. \tag{3.14}$$

In this model the relative valuation V is a complicated function of all parameters and endogenous variables. We summarize the properties of $\partial V/\partial\kappa$ and $\partial V/\partial\kappa_D$ in the following proposition

Proposition 6. *Assume that* $\gamma > 1$, *then*

$$\frac{\partial V}{\partial \kappa} > 0$$

and, for all γ,

$$\frac{\partial V}{\partial \kappa_D} = 0 \Leftrightarrow 1 - V = (1-\gamma)\frac{\partial P_D/\partial \kappa_D - 1}{\upsilon'(\kappa_D)}.$$

Proof. (see the Appendix) □

Given that $\partial V/\partial \kappa > 0$ when $\gamma > 1$, equation (3.13) implies that

$$\frac{d\mu^{\delta}}{d\kappa}(\kappa) < \frac{1}{\eta},$$

which shows that optimal prevention in this case requires more investment than what would be required to equate the marginal cost of prevention —which is one in this case— with the marginal benefit of reducing the losses of stocks —which in this case is $\eta d\mu^{\delta}/d\kappa$. The reason is simple: In this specification investments in prevention have a positive impact on flow productivity of both forms of capital during the disaster phase. Hence this second component increases the marginal benefit.

The optimal level of mitigation implies that

$$z(\kappa_D) \equiv \frac{\partial P_D/\partial \kappa_D - 1}{\upsilon'(\kappa_D)}$$

must satisfy

$$z(\kappa_D) = \frac{1-V}{1-\gamma}. \tag{3.15}$$

Given the results of Proposition 5 the right hand side of equation (3.15) is negative and this implies that, at the optimum, the marginal product of mitigation investments is less than the marginal cost. The reason is simple: mitigation also shortens the expected duration of the disaster phase and this a valuable

$$\partial P_D/\partial \kappa_D < 1.$$

Under some assumptions about the specific technologies the function $z(\kappa_D)$ is downward sloping. In that case the results in Proposition 5 imply that increases in $\bar{\zeta}$ decrease the optimal κ_D. Thus, higher expected foreign aid weakens the incentives that the country has to invest in activities that increase productivity and shorten the duration of the disaster phase. In particular, the model implies that countries that receive a higher level of foreign aid in response to a natural disaster strike will experience longer periods of low productivity.

3.3 Quantitative Analysis

In this section we report the results from parameterizing the model. At this point our quantitative exercise is aimed at trying to understand the interplay between different mechanisms and forces in the model rather than trying to match any country's experience. Moreover, we were not able to find reliable data in order to estimate the relevant parameters. Instead we report the criteria that we used to select specific values.

3.3.1 Functional Forms

As indicated in the model the productivity in the disaster phase is lower than in the normal phase. We assume the post-strike productivities follow

$$\begin{aligned} A_D &= A f_A(\kappa, \kappa_D) \\ H_D &= H f_H(\kappa, \kappa_D) \end{aligned}$$

with

$$f_j \in (0,1)\,; \quad j = \{A, H\}$$

and

$$\begin{aligned}\frac{\partial f_j}{\partial \kappa} &> 0 \\ \frac{\partial f_j}{\partial \kappa_D} &> 0\end{aligned}$$

More specifically, we specify that the physical and human productivities in the disaster phase are given by

$$\begin{aligned} f_A(\kappa, \kappa_D) &= 1 - \phi_A e^{-(\lambda^A \kappa + \lambda_D^A \kappa_D)} \\ f_H(\kappa, \kappa_D) &= 1 - \phi_H e^{-(\lambda^H \kappa + \lambda_D^H \kappa_D)} \end{aligned}$$

where ϕ_A and ϕ_H are the respective productivity losses under zero investment in prevention and mitigation and λ^A and λ_D^A (λ^H and λ_D^H) are the semi-elasticities of the physical (human) capital stock loss functions (i.e. $-\phi_j e^{-(\lambda^j \kappa + \lambda_D^j \kappa_D)}$) with respect to investment in prevention and mitigation, respectively.

We assume that the probability of returning to the *normal* phase after the disaster hits the economy is given by

$$\upsilon(\kappa_D) = \upsilon_0 (1 + \upsilon_1 \kappa_D)^{\upsilon_2}$$

where υ_0 is the inverse of the expected disaster duration under no investment in mitigation, and υ_1 and υ_2 are scale and curvature parameters.

The losses of the two stocks when the natural disaster hits depend on the amount of prevention resources, κ, according to

$$\mu_k^{\delta}(\kappa) = 1 - \mu_k^0 e^{-\mu_k^1 \kappa}$$
$$\mu_h^{\delta}(\kappa) = 1 - \mu_h^0 e^{-\mu_h^1 \kappa}$$

where μ_k^0 and μ_h^0 are the physical and human capital stock losses under zero prevention investment and μ_k^1 and μ_h^1 are the semi-elasticities of physical and human capital stock losses with respect to prevention.

Our aim in choosing these functional forms was to present a fairly general approach to trying to capture the role of prevention and mitigation in reducing the impact of a natural disaster on the productive capabilities of an economy.

3.3.2 Calibration

Given the limited data availability on natural disasters and their impact, we take what we consider a reasonable calibration and we analyze the sensitivity of the results obtained to changes to this baseline case.

We assume that in the event of natural disaster, and under zero investment in prevention and mitigation, the physical and human capital productivities fall by 20% and 10%, respectively. Thus $\phi_A = 0.2$ and $\phi_H = 0.1$. Furthermore, we initially consider the case of equal impact of prevention and mitigation on physical and human capital productivities, which implies $\lambda^A = \lambda_D^A$ and $\lambda^H = \lambda_D^H$. To pin down λ^A and λ^H we take the approach that at relatively high levels of investment in prevention and mitigation [9] their marginal impact on

[9] We think of $\kappa = 0.05$ and $\kappa_D = 0.05$ as those levels, which, considering a total wealth -physical and human capital- to output ratio of six, amount to prevention and mitigation investments of around 30% of GDP.

productivity is almost negligible. Thus, λ^A and λ^H solve

$$\begin{aligned}
\lambda^A 0.2 e^{-(\lambda^A \kappa + \lambda_D^A \kappa_D)} \quad & |_{\kappa=0.05;\kappa_D=0.05} \approx 0 \\
\lambda^H 0.1 e^{-(\lambda^H \kappa + \lambda_D^H \kappa_D)} \quad & |_{\kappa=0.05;\kappa_D=0.05} \approx 0.
\end{aligned}$$

We find that $\lambda^A = \lambda_D^A \approx 140$ and $\lambda^H = \lambda_D^H \approx 110$.

In the case of the stock loss functions we suppose that without investment in prevention the physical and human capital losses are 10% and 5% of their corresponding stocks, respectively

$$\begin{aligned}
\mu_k^0 &= 0.10 \\
\mu_h^0 &= 0.05
\end{aligned}$$

We calibrate μ_k^1 and μ_h^1 in a similar fashion as the productivity loss functions, considering a nearly zero impact of additional investment in prevention for high investment levels. In this case μ_k^1 and μ_h^1 solve

$$\begin{aligned}
0.1 \mu_k^1 e^{-\mu_k^1 \kappa} \quad & |_{\kappa=0.05} \approx 0 \\
0.05 \mu_h^1 e^{-\mu_h^1 \kappa} \quad & |_{\kappa=0.05} \approx 0
\end{aligned}$$

which result in $\mu_k^1 \approx 145$ and $\mu_h^1 \approx 130$.

For the disaster recovery probability, we calibrate υ_0 such that the expected duration of the disaster phase is three years under zero investment in mitigation, hence $\upsilon_0 = \frac{1}{3}$. We choose υ_1 and υ_2 such that the expected recovery speed is 1.5 years and 0.5 years for prevention investments of 5% and 30% of GDP, respectively. Hence, υ_1 and υ_2 solve

$$\begin{aligned} 1.5 &= \frac{1}{3}(1+\upsilon_1 \frac{0.05}{6})^{\upsilon_2} \\ 0.5 &= \frac{1}{3}(1+\upsilon_1 \frac{0.30}{6})^{\upsilon_2} \end{aligned}$$

We obtain $\upsilon_1 \approx 510$ and $\upsilon_2 \approx 0.55$.

We take the discount rate to be $\rho = 0.04$, the relative risk aversion parameter $\gamma = 2$, and the probability of a natural disaster hitting the economy to be 0.03. The latter implies that a natural disaster hits the economy every 33 years on average, a very rare event. We take the expected return on physical capital to be 10% ($A = 0.1$), the return on human capital to be 6% ($H = 0.06$) and we calibrate the volatility of physical capital to match the historical return volatility of the S&P 500 ($\sigma = 0.16$).

3.3.3 Sensitivity: Quantitative Results

In this section we analyze the model's quantitative behavior under the baseline calibration, starting with the no disaster insurance case. We set the generic wealth (ζ_w) and stock-specific transfers (ζ_k and ζ_h) and find the utility maximizing levels of $(\kappa, \kappa_D, \alpha_N, \alpha_D)$.[10] We report those values as well as the growth rates on the two phases, μ_N and μ_D, the fraction of total wealth left over after a disaster strikes, $\beta(\alpha, \kappa, \mu^\delta, \zeta)$, and the expected duration of the disaster phase, $1/\upsilon(\kappa_D)$.

We label our base scenario, *Case 1.* We then explore the sensitivity of the endogenous choices to changes in the basic parameterization that we label Cases 2-6.

- *Case 2*: This case displays higher semi-elasticities of investment of both prevention

[10] We search over a grid of 61 equally spaced values for κ and κ_D between 0 and 0.05 to find the investment levels in prevention and mitigation that maximize utility.

and mitigation for both forms of capital. We choose $\lambda_D^A = 2\lambda^A$ and $\lambda_D^H = 2\lambda^H$.

- *Case 3:* We allow for less curvature (higher marginal product) of investments in prevention and mitigation. We capture this by requiring that the marginal product be low (close to zero) when $\kappa = \kappa_D = 0.1$, instead of 0.05.

- *Case 4:* We increase the losses of the two stocks when a natural disaster increases. If no efforts in prevention and mitigation are undertaken we assume that 40% of the physical capital stock is lost and 20% of the human capital is destroyed. Thus, $\phi'_A = 0.4$ and $\phi'_H = 0.2$

- *Case 5:* We triple the volatility of the risky technology and increase σ from 0.16 to 0.48.

- *Case 6:* We assume that a natural disaster occurs, on average, every ten years. This changes η from 0.03 to 0.1.

The results are reported in Table 1 below

Table 3.1: Sensitivity to Changes in the Model's Parametrization

-	*Case 1*	*Case 2*	*Case 3*	*Case 4*	*Case 5*	*Case 6*
κ	0.00%	0.00%	0.00%	0.00%	0.00%	0.42%
κ_D	1.20%	1.00%	1.50%	1.70%	0.60%	0.42%
μ_N	3.37%	3.37%	3.38%	3.37%	1.35%	3.20%
μ_D	1.74%	2.25%	0.86%	1.26%	0.39%	2.43%
α_N	0.746	0.746	0.745	0.745	0.08	0.726
α_D	0.701	0.735	0.638	0.706	0.07	0.651
$\frac{1}{\upsilon(\kappa_D)}$	1.03	1.11	0.92	0.87	1.40	1.60
$\beta(\alpha, \kappa, \mu^\delta, \zeta)$	0.91	0.91	0.91	0.91	0.95	0.95

The quantitative results do not vary significantly between cases *Cases 1 to 4*. In all cases it is optimal to spend no resources in prevention. Rather it is better to spend between 6% and 10% of GDP in the disaster phase increasing productivity and shortening the duration of that phase. The highest level of spending in mitigation occurs when the natural disaster is costliest in terms of stock losses (*Case 4*)

We find that the growth rate of output displays little change across all four cases in the normal phase, while the growth rate in the disaster phase depends on the impact of the disaster on productivity and the marginal returns of the mitigation technology. The intuition is straightforward: since a disaster is a very unlikely event and mitigation measures can be adopted instantaneously, it is optimal to save on prevention resources, keep a high growth rate in the normal phase and take remedy measures when the economy is hit by a disaster.

The main difference in the first four cases is the resulting growth rate in the disaster phase. In *Case 2*,the higher marginal return on mitigation allows for a reduction in mitigation investment —relative to *Case 1* -while still keeping a higher post-strike productivity ($A_D^2 = 0.099 > A_D^1 = 0.096$). This drives up the portfolio share of the risky asset. These two elements combined result in a higher growth rate in the disaster phase. *Case 3* in, in one dimension, the opposite of *Case 2.* Given our calibration, it implies that the marginal impact of prevention and mitigation are smaller for similar investment levels than in *Case 1.*The mechanism driving the results is, therefore, the same as in *Case 2,* but acting in the opposite direction.

Case 4 requires a higher mitigation investment to keep the return of the risky asset at the same level as in *Case 1.* As a result, even if the capital portfolio shares are very close in both cases, the higher required mitigation investment drags growth down in the disaster phase.

In *Case 5* (higher volatility of the risky technology) we find, as expected, a sizable decrease in the fraction of wealth allocated to the risky asset, both in the normal and in the disaster phase. This drives down growth under both scenarios. Relative to the previous four cases it also reduces the efforts at mitigation.

As noted above, a common feature of *Cases 1* to *5* is that the optimal investment on prevention is zero in all of them. This is driven by our assumption that natural disasters are rare. When we increase the expected arrival time from 33 to 10 years we find a positive investment in prevention. We view this case as somewhere "in between" a parameterization that applies to earthquakes —fairly rare events— and a parameterization that captures the impact of hurricanes —an almost yearly occurrence.

3.3.4 Aid and Welfare: Growth Effects

In this section we explore the interplay between foreign aid and the endogenous choice of investments in prevention and mitigation. We consider three different scenarios in terms of how rare natural disasters are. The three cases area natural disaster on average every 33 years ($\eta = 0.03$). every 10 years ($\eta = 0.10$), and every two years ($\eta = 0.50$).

We initially set the foreign aid (transfers) equal to zero and increase them up to 10% of the post-strike wealth. In this preliminary exercise we only study the impact of "general" (as opposed to stock specific) transfers. This corresponds to what we labeled ζ_w in the theoretical model.

In Tables 2-4 we report the same variables as in Table 1. In addition we indicate the welfare gain —relative to the no transfer case— associated with foreign aid. We follow standard practice in macro and estimate the welfare gains as the percentage increase in permanent consumption associated with the transfer.[11] Of course, higher foreign aid increases welfare in a monotonic way but transfers have less obvious effects:

1. In all cases higher transfers lower the growth rate in both phases. This decrease is driven by the country adjusting the level of overall saving, the composition of the portfolio and the investments in prevention and mitigation.

[11] Label $\overline{C}(\zeta) = \overline{c}w$ the constant level of lifetime consumption that yields utility $V_N(w)$. Thus $V_N(\zeta)\frac{w^{1-\gamma}}{1-\gamma} \equiv \frac{(\overline{c}(\zeta)w)^{1-\gamma}}{1-\gamma}\frac{1}{\rho}$, which implies $\overline{c}(\zeta) = (\rho V_N(\zeta))^{\left(\frac{1}{1-\gamma}\right)}$ and $\overline{c}(0) = (\rho V_N(0))^{\left(\frac{1}{1-\gamma}\right)}$ .Therefore $\frac{\overline{c}(\zeta)}{\overline{c}(0)} = \left[\frac{V_N(0)}{V_N(\zeta)}\right]^{\left(\frac{1}{1-\gamma}\right)}$.

2. In the rare event scenario ($\eta = 0.03$) foreign aid and investment in mitigation are complements: the higher the transfer the higher the optimal investment in mitigation. Moreover, total saving is decreasing in the transfer in the normal phase and almost constant in the disaster phase. The portfolio effects of foreign aid are small.

3. In the intermediate and frequent case scenarios ($\eta = 0.10$ and $\eta = 0.50$) foreign aid and investment in prevention are substitutes but foreign aid and investment in mitigation are complements. The impact of aid on total investment is positive: investment in mitigation increases in a magnitude than more than compensates the fall in investment in prevention. The portfolio effects are small.

4. Growth and welfare move in opposite directions: the higher the level of foreign aid the higher the welfare and the lower the growth rate in both phases.

5. The growth impact of foreign aid is significantly smaller (although still negative) in the normal phase. Since higher transfers are associated with lower investment in prevention it must be the case that, in the normal phase, the expectation of higher transfers lowers saving in both productive assets.

6. Growth reversals. The model implies that in the case of rare natural disasters the growth rate in the normal phase is higher than in the disaster phase. However, for events that, on expectation, happen every two years the opposite is true: growth is higher in the disaster phase.

Table 3.2: Growth and Welfare Effects of Foreign Aid
(*No Disaster Insurance -*$\eta = 0.03$*-*)

ζ_w	**0.00**	**0.01**	**0.02**	**0.05**	**0.10**
κ	0.000	0.000	0.000	0.000	0.000
κ_D	0.0117	0.0117	0.0117	0.0125	0.0133
μ_N	3.37%	3.36%	3.34%	3.30%	3.24%
μ_D	1.74%	1.72%	1.63%	1.63%	1.53%
α_N	0.746	0.746	0.747	0.749	0.752
α_D	0.701	0.701	0.701	0.707	0.712
$\frac{1}{\upsilon(\kappa_D)}$	1.033	1.033	1.033	1.000	0.969
c_N	0.742	0.745	0.747	0.754	0.764
c_D	0.772	0.774	0.776	0.779	0.785
$\frac{\kappa}{F_N}$	0.000	0.000	0.000	0.000	0.000
$\frac{\kappa_D}{F_D}$	0.162	0.162	0.162	0.172	0.183
$\beta(\alpha, \kappa, \mu^\delta, \zeta)$	0.91	0.92	0.93	0.99	1.01
$\Delta Welfare$	0.00	0.60%	1.40%	3.00%	5.80%

Table 3.3: Growth and Welfare Effects of Foreign Aid
(*No Disaster Insurance -*$\eta = 0.10$*-*)

ζ_w	**0.00**	**0.01**	**0.02**	**0.05**	**0.10**
κ	0.0042	0.003	0.003	0.0025	0.0017
κ_D	0.0042	0.007	0.007	0.0100	0.0142
μ_N	3.20%	3.21%	3.16%	3.08%	2.92%
μ_D	2.43%	2.27%	2.22%	1.91%	1.40%
α_N	0.726	0.719	0.720	0.714	0.710
α_D	0.651	0.660	0.660	0.669	0.673
$\frac{1}{\upsilon(\kappa_D)}$	1.60	1.33	1.33	1.11	0.94
$\beta(\alpha, \kappa, \mu^\delta, \zeta)$	0.87	0.87	0.87	0.92	0.91
$\frac{c_N}{Y_N}$	0.700	0.706	0.713	0.732	0.762
$\frac{c_D}{Y_D}$	0.750	0.745	0.751	0.757	0.774
$\frac{\kappa}{Y_N}$	0.055	0.044	0.044	0.033	0.022
$\frac{\kappa_D}{Y_D}$	0.059	0.094	0.094	0.138	0.193
$\Delta Welfare$	0	1.85%	3.70%	9.23%	18.3%

Table 3.4: Growth and Welfare Effects of Foreign Aid
(*No Disaster Insurance* -$\eta = 0.50$-)

ζ_w	**0.00**	**0.01**	**0.02**	**0.05**	**0.10**
κ	0.020	0.020	0.020	0.020	0.0142
κ_D	0.000	0.000	0.000	0.000	0.0175
μ_N	1.88%	1.77%	1.66%	1.35%	1.13%
μ_D	3.66%	3.57%	3.48%	3.21%	0.84%
α_N	0.765	0.765	0.766	0.766	0.741
α_D	0.753	0.753	0.753	0.753	0.733
$\frac{1}{\upsilon(\kappa_D)}$	3.00	3.00	3.00	3.00	0.849
$\beta(\alpha, \kappa, \mu^\delta, \zeta)$	0.99	1.00	1.01	1.05	1.09
$\frac{c_N}{Y_N}$	0.686	0.700	0.714	0.756	0.849
$\frac{c_D}{Y_D}$	0.703	0.716	0.728	0.765	0.839
$\frac{\kappa}{Y_N}$	0.264	0.264	0.264	0.264	0.187
$\frac{\kappa_D}{Y_D}$	0.000	0.000	0.000	0.000	0.232
$\Delta Welfare$	0	4.20%	8.50%	21.5%	53.0%

3.4 Conclusion

We develop a continuous time stochastic growth model that is suitable for studying the impact of natural disasters on the short run and long run growth rate of an economy.Even though the Poisson shocks that we use to capture the large and unusual natural disaster shocks are not of the more standard variety, we find that the growth impact of shocks depends on the curvature of the utility function. Additionally, we find that the growth effects of a natural disaster depend in complicated ways on the existence of catastrophe insurance markets. Interestingly, our results show that foreign aid received when the natural disaster impacts a country in the normal phase and aid when the country is in the disaster regime have potentially opposite impacts on the growth rate during the recovery period. Moreover, under reasonable conditions on the prevention and mitigation technologies, increases in foreign aid reduce investment in mitigation activities and, as a result, delay the recovery from the disaster.

This work constitutes a first attempt to understand the effects of foreign aid and disaster insurance on the growth rate of the economy hit by such event, not only upon arrival but also in the recovery phase. Given the importance of the presence of insurance markets on the incentives that countries face, future work should address optimal design of such markets.

Bibliography

Acemoglu, D. (1998): "Why Do New Technologies Complement Skills? Directed Technical Change and Wage Inequality," *Quarterly Journal of Economics*, 113, 1055–89.

——— (2002): "Directed Technical Change," *Review of Economic Studies*, 69, 781–809.

Acemoglu, D. and D. Autor (2011a): *Skills, Tasks, and Technologies: Implications for Employment and Earnings*, North-Holland: Elsevier.

——— (2011b): "Skills, Tasks and Technologies: Implications for Employment and Earnings," Elsevier, vol. 4 of *Handbook of Labor Economics*, chap. 12, 1043–1171.

Acemoglu, D., F. Gallego, and J. A. Robinson (2014): "Institutions, Human Capital, and Development," *The Annual Review of Economics*, 6, 875–912.

Acemoglu, D. and V. Guerrieri (2008): "Capital Deepening and Nonbalanced Economic Growth," *Journal of Political Economy*, 116, 467–498.

Autor, D. H. and D. Dorn (2013): "The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market," *American Economic Review*, 103, 1553–97.

Autor, D. H., D. Dorn, and G. H. Hanson (2013): "The Geography of Trade and Technology Shocks in the United States," *American Economic Review*, 103, 220–25.

Autor, D. H., L. F. Katz, , and A. B. Krueger (1998): "Computing Inequality: Have Computers Changed the Labor Market?" *Quarterly Journal of Economics*, 113, 1169–1213.

Autor, D. H., L. F. Katz, and M. S. Kearney (2008): "Trends in U.S. Wage Inequality: Revising the Revisionists," *The Review of Economics and Statistics*, 90, 300–323.

Bakkensen, L. and L. Barrage (2017): "Do Disasters Affect Growth? A Macro-Model Based Perspective on the Empirical Debate," *Working Paper*.

Barro, R. (2009): "Rare Disasters, Asset Prices and Welfare Costs," *American Economic Review*, 99, 243–264.

Baumol, W. J. (1967): "Macroeconomics of Unbalanced Growth: Anatomy of an Urban Crisis," *American Economic Review*, 57, 415–26.

Berlemann, M. and D. Wenzel (2016): "Hurricanes, Economic Growth and Transmission Channels," *CESifo Working Paper No. 6041*, -, –.

Berman, E., J. Bound, and Z. Griliches (1994): "Changes in the Demand for Skilled Labor within U.S. Manufacturing: Evidence from the Annual Survey of Manufactures," *The Quarterly Journal of Economics*, 109, 367–97.

Berman, E., J. Bound, and S. Machin (1998): "Implications Of Skill-Biased Technological Change: International Evidence," *The Quarterly Journal of Economics*, 113, 1245–1279.

Bick, A., N. Fuchs-Schundeln, and D. Lagakos (2018): "How Do Hours Worked Vary with Income? Cross-Country Evidence and Implications," *American Economic Review*, 108, 170–199.

Boppart, T. (2014): “Structural Change and the Kaldor Facts in a Growth Model With Relative Price Effects and Non?Gorman Preferences,” *Econometrica*, 82, 2167–2196.

Bound, J. and G. Johnson (1992): “Changes in the Structure of Wages in the 1980’s: An Evaluation of Alternative Explanations,” *American Economic Review*, 82, 371–92.

Buera, F., R. Rogerson, J. Kaboski, and J. I. Vizcaino (2018): “Skill-Biased Technical Change,” *Manuscript*.

Buera, F. J. and J. P. Kaboski (2009): “Can Traditional Theories of Structural Change Fit The Data?” *Journal of the European Economic Association*, 7, 469–477.

——— (2012): “The Rise of the Service Economy,” *American Economic Review*, 102, 2540–69.

Caselli, F. (2005): *Accounting for Cross-Country Income Differences*, North-Holland: Elsevier.

——— (2016): “Technology Differences Over Space and Time,” *Manuscript*, -, –.

Caselli, F. and A. Ciccone (2019): “The Human Capital Stock: A Generalized Approach: Comment,” *American Economic Review*, 109, 1155–74.

Caselli, F. and W. J. I. Coleman (2006): “The World Technology Frontier,” *American Economic Review*, 96, 500–522.

Comin, D., D. Lakshari, and M. Mestieri (2015): “Structural Change with Long Run Income and Price Effects,” Manuscript.

Deryugina, T. (2017): “The Fiscal Cost of Hurricanes: Disaster Aid Versus Social Insurance,” *American Economic Journal: Economic Policy*, 9, 168–198.

DiNardo, J., N. M. Fortin, and T. Lemieux (1996): "Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach," *Econometrica*, 64, 1001–44.

Erosa, A., T. A. Koreshkova, and D. Restuccia (2010): "How important is Human Capital: A Quantitative Theory Assesment of World Income Inequality," *Review of Economic Studies*, 52, 1–32.

Feenstra, R. C. and G. H. Hanson (1999): "The Impact Of Outsourcing And High-Technology Capital On Wages: Estimates For The United States, 1979-1990," *The Quarterly Journal of Economics*, 114, 907–940.

Feenstra, R. C., R. Inklaar, and M. P. Timmer (2015): "The Next Generation of the Penn World Table," *American Economic Review*, 105, 3150–3182.

Gabaix, X. (2012): "Variable Rare Disasters: An Exactly Solved Framework for Ten Puzzles in Macro-Finance," *The Quarterly Journal of Economics*, 127, 645–700.

Gollin, D., D. Lagakos, and M. E. Waugh (May 2014): "The Agricultural Productivity Gap," *The Quarterly Journal of Economics*, 129, 939–993.

Gourio, F. (2012): "Disaster Risk and Business Cycles," *American Economic Review*, 102, 2734–2766.

Hall, R. E. and C. I. Jones (1999): "Why Do Some Countries Produce So Much More Output Per Worker Than Others?" *The Quarterly Journal of Economics*, 114, 83–166.

——— (2007): "The Value of Life and the Rise in Health Spending," *The Quarterly Journal of Economics*, 122, 39–72.

Heckman, J. T. and B. E. Honoré (1990): "The Empirical Content of The Roy Model," *Econometrica*, 58, 1121–49.

Herrendorf, B., R. Rogerson, and A. Valentinyi (2013): “Two Perspectives on Preferences and Structural Transformation,” *American Economic Review*, 103, 2752–89.

——— (2014): “Growth and Structural Transformation,” in *Handbook of Economic Growth*, Elsevier, vol. 2 of *Handbook of Economic Growth*, chap. 6, 855–941.

Hsiang, S. and A. Jina (2014): “The Causal Effect of Environmental Catastrophe on Long-Run Economic Growth: Evidence from 6,700 Cyclones,” *NBER Working Paper 20352*, -, –.

Jaume, D. (2019): “The Labor Market Effects of an Educational Expansion. A Theoretical Model with Applications to Brazil,” *Manuscript*, -, –.

Jones, B. and B. Olken (2008): “The Anatomy of Start-Stop Growth,” *The Review of Economics and Statistics*, 87, 271–284.

Jones, B. F. (November 2014): “The Human Capital Stock: A Generalized Approach,” *American Economic Review*, 104, 3752–77.

Katz, L. F. and D. H. Autor (1999): *Changes in the wage structure and earnings inequality*, North-Holland: Elsevier.

Katz, L. F. and K. M. Murphy (1992a): “Changes in Relative Wages, 1963-1987: Supply and Demand Factors,” *Quarterly Journal of Economics*, 107, 35–78.

——— (1992b): “Changes in Relative Wages, 1963-1987: Supply and Demand Factors,” *The Quarterly Journal of Economics*, 107, 35–78.

Klenow, P. J. and A. Rodriguez-Clare (1997): “The Neoclassical Revival in Growth Economics: Has It Gone Too Far?” *in: B.S. Bernanke, and J.J. Rotemberg, eds., NBER 66 macroeconomics annual.*

Kongsamut, P., S. Rebelo, and D. Xie (2001): "Beyond Balanced Growth," *Review of Economic Studies*, 68, 869–82.

Kousky, C. (2014): "Informing Climate Adaptation: A Review of the Economic Costs of Natural Disasters," *Energy Economics*, 46, 576–592.

Krusell, P., L. E. Ohanian, J.-V. Ros-Rull, and G. L. Violante (2000): "Capital-Skill Complementarity and Inequality: A Macroeconomic Analysis," *Econometrica*, 68, 1029–1054.

Lagakos, D. and M. E. Waugh (April 2013): "Selection, Agriculture, and Cross-Country Productivity Differences," *American Economic Review*, 103, 948–980.

Leonardi, M. (forthcoming): "The Effect of Product Demand on Inequality: Evidence from the US and the UK," *American Economic Journal: Applied*.

Malmberg, H. (2018): "How does the eciency of skilled labor vary across rich and poor countries? An analysis using trade and industry data,," *Manuscript*, -, –.

Mankiw, G. M., D. Romer, and D. N. Weil (1992): "A Contribution To The Empirics of Economic Growth," *The Quarterly Journal of Economics*, 107, 407–437.

Manuelli, R. E. and A. Seshadri (September 2014): "Human Capital and the Wealth of Nations," *American Economic Review*, 104.

Murphy, K. M. and F. Welch (1992): "The Structure of Wages," *The Quarterly Journal of Economics*, 107, 285–326.

Nelsen, R. B. (2006): *An Introduction to Copulas*, Springer.

Ngai, L. R. and B. Petrongolo (2014): "Gender gaps and the rise of the service economy," LSE Research Online Documents on Economics 58152, London School of Economics and Political Science.

Ngai, L. R. and C. A. Pissarides (2007): "Structural Change in a Multisector Model of Growth," *American Economic Review*, 97, 429–443.

O'Mahony, M. and M. P. Timmer (2009): "Output, Input and Productivity Measures at the Industry Level: The EU KLEMS Database," *Economic Journal*, 119, 374–403.

Pierce, B. (2010): "Recent Trends in Compensation Inequality," in *Labor in the New Economy*, National Bureau of Economic Research, Inc, 63–98.

Reshef, A. (2013): "Is Technological Change Biased Towards the Unskilled in Services? An Empirical Investigation," *Review of Economic Dynamics*, 16, 312–331.

Rossi, F. (2017): "The Relative Efficiency of Skilled Labor across Countries: Measurement and Interpretation," *Manuscript*, -, –.

Skidmore, M. and T. Toya (2007): "Do Natural Disasters Promote Long-Run Growth?" *Economic Inquiry*, 40, 664–687.

Swiecki, T. (2014): "Determinants of Structural Change," Manuscript, University of British Columbia.

von Peter, G., S. von Dahlen, and S. Saxena (2012): "Unmitigated Disasters? New Evidence on the Macroeconomic Cost of Natural Catastrophes," *BIS working paper No. 394*, -, –.

Appendix A

Supplementary Material to Chapter 1

A.1 ILO Data Description, Treatment, and Issues.

The ILO's database is constructed from multiple data sources, including establishment surveys, household surveys, insurance records, and administrative data sources. Data sources vary across countries, and when multiple sources are available, ILO presents all the options available. In such case, I pick data from the source I consider most reliable [1]. I discard data from sources that the ILO flags as unreliable, even if that is the only one available for a country.

The ILO data is harmonized, both at the sectoral and occupational level, allowing for comparability across countries. Occupational data is harmonized based on the International Standard Classification of Occupations (ISCO). Statistics on employment by occupation are presented in ILOSTAT according to both the categories of the latest version of the ISCO available (ISCO-08 and ISCO-88). When both versions are available, I take the latest revision (ISCO-08). I take the earlier version (ISCO-88) when it is the only one available and bridge it into the newer (ISCO-08) using the correspondence table provided by ILO [2].

A.2 Countries Intensive in Natural Resources

Natural resources rents are measured by the World Bank's World Development Indicators variable Total natural resources rents (% of GDP)(NY.GDP.TOTL.RT.ZS). Total natural resources rents are the sum of oil rents, natural gas rents, coal rents (hard and soft), mineral rents, and forest rents.

[1] Each source has it's own advantages and disadvantages, depending on the country under study. For example, establishment data tend to be very accurate, but it has limitations in countries where firms routinely pay wages outside their normal book-keeping in order to avoid taxes. Household surveys cover all employees regardless of where they work, but their reliability depends heavily on the accuracy of the respondent.

[2] Link to ILO's occupations documentation.

To decide which countries to exclude I take the 217 countries with data available in the WDI database and rank them according to their average natural resources rents as percentage of GDP for the period 2008-2012. The criterion for exclusion is to discard the countries in the top decile in terms of natural resources rents. I end up excluding 21 countries with natural resources rents above 25.1% of GDP.

The countries excluded are Libya (53.2%), Kuwait (52.4%), Republic of Congo (51.4%), Saudi Arabia (46.3%), Iraq (45.2%), Mauritania (44.1%), Angola (42.3%), Oman (40.6%), Papua New Guinea (39.9%), Liberia (39.8%), South Sudan (38.4%), Gabon (37.4%), Equatorial Guinea (35.3%), Mongolia (34.1%), Turkmenistan (33.1%), Azerbaijan (32.5%), Chad (28.4%), Guinea (26.7%), Burundi (25.9%), and Brunei Darussalam (25.1%).

A.3 Hours Worked Across Development.

In this section I briefly analyze the evolution of hours worked as countries develop. It is of special interest to see how my data compares to that in Bick, Fuchs-Schundeln, and Lagakos Bick et al. (2018), the paper that, up to my knowledge, more carefully studies this issue. Data limitations allow me to compare hours per worker by sector only. Even though the authors do not study the evolution of hours worked by occupation, I show how hours worked differ for my broad occupational categories across development as well.

The Table above shows that although average hours worked follow the same pattern in both

Table A.1: Average Hours Worked by Sector Across Development
-ILO data vs B,F-S,L-

2*Income Group	**Agriculture**		**Industry**		**Services**	
	ILO data	*B,F-S,L*	*ILO data*	*B,F-S,L*	*ILO data*	*B,F-S,L*
Low	35.0	36.0	45.7	44.9	46.0	47.7
Middle	38.5	38.3	42.8	42.5	41.8	41.8
High	41.5	39.7	39.6	37.0	37.2	34.7

Note I : B,F-S,L corresponds to the data in Bick, Fuchs-Schundeln, and Lagakos. In their paper Industry is called Manufacturing. The sectors included in it are roughly the same except for Construction, which I include in Industry and it is not clear to me if they consider it to be in Manufacturing or Services.
Note II : To compute GDP per capita and development percentiles I use as GDP measure PWT's *rgdpo*. I take the average GDP per capita for the period 2005-2014. B,F-S,L use as GDP measure PWT's *rgdpe*, and compute terciles for 2005. I follow their procedure to construct this table.
Note III : B,F-S,L focus on 49 core countries while I have data for 88 countries.

cases, magnitudes differ. For instance,in Agriculture there is a 6.5 hour increase between the top and the bottom development tercile, while this number falls to 3.9 in B,F-S,L. At the same time, the decline in hours worked in Industry (6.1 vs 7.9) and Services (8.8 vs 13.0) as countries move from the bottom to the top tercile is smaller in my sample as compared to B,F-S,L. The second Table in this section shows that the pattern described in B,F-S, L is robust to using different measures of GDP per capita as a proxy of development and to splitting to sample into quartiles instead of terciles. Interestingly, the decline in average hours worked is more pronounced in low-skill than in high-skill services (-11.1 vs -7.2 if one

Table A.2: Average Hours Worked by Sector Across Development
-using ILO data and four broad sectors-

Income Quartile	**Countries**	**Agriculture**	**Industry**	**L-S Services**	**H-S Services**
Bottom	22	35.6	46.1	46.8	46.0
2nd	18	36.6	44.0	44.6	42.3
3rd	25	38.4	41.8	40.5	40.6
Top	23	42.2	39.6	35.7	38.8

Note I : To compute income quartiles I consider all countries in PWT 9.0, use as GDP measure PWT's *rgdpo* and take the average GDP per capita for the period 2005-2014. This criterion differs from B,F-S,L, as described in Note II of the Table above.

compares the Top and the Bottom quartiles).

Table A.3: Average Hours Worked by Occupation Across Development
-using ILO data-

Income Quartile	**high-skill**	**Low-Skill**
Bottom	40.9	46.0
2nd	40.8	47.4
3rd	40.0	43.1
Top	37.6	37.1

A.4 Robustness Checks.

I here explain in further detail the robustness checks I perform to my quantitative empirical analysis in Section 1.3.

Different Criteria for Constructing Broad Occupational Groups.

I proceed to study if the results presented in Section 1.3.1 depend on the criterion used to group one-digit ISCO-08 categories into high- and low-skill occupations. I perform a sensitivity analysis under three different grouping criteria.

The first exercise moves the occupational group with lowest median wages at all development levels in the high-skill group, namely Clerical, from the high-skill to the low-skill group. The second exercise, moves one of the categories with highest median wages in low-skill occupations, Service Workers, into the high-skill group. It is worth pointing out that now the original grouping criterion is no longer respected, as median wages for Service Workers are, in this case, lower than those for Skilled Agricultural Workers in the second and third development quartiles. However, I consider that switching one occupational group at a time presents a more thorough and transparent way to asses how my results are affected by different grouping criteria. In my third exercise I include both Service and Skilled Agricultural Workers in the high-skill group. In this case it holds that median wages in for all the occupations in the high-skill group are higher than those in the low-skill group, at all development levels [3]

The results are presented in Table A.4 below. Compared to the baseline case presented in Column 1 of Table A.4, switching occupational group four from high- to low-skill occupations has the effect of increasing both the constant and the skill-premium elasticities for all

[3] The same order holds when I compute average wages by development quartiles. I prefer classify my occupations using median instead of average wages by quartile because the former measure does not depend on extreme values, a feature that is particularly present in my data at low development levels.

quartiles. Estimated coefficients are both individually and jointly statistically significant at the one percent level.

Switching either occupation seven or seven and eight together from the Low- to the high-skill groups have similar effects. The regression constant falls, more in the latter case, while the skill-premium elasticities are in both instances smaller compared to the baseline, but roughly speaking, very similar. In these two counterfactuals the estimated coefficients are jointly significant at the one percent level, while individual coefficients are now significant at the ten percent level, at least.

Table A.4: Development Elasticity of the Occupational Skill-Premium
Sensitivity to Different Grouping Criteria

	(1)	(2)	(3)	(4)
	Baseline	-	-	-
	High-Skill	High-Skill	High-Skill	High-Skill
	1,2,3,4	1,2,3	1,2,3,4,7	1,2,3,4,7,8
Variables	$log\left(\frac{w_{hs}}{w_{ls}}\right)$	$log\left(\frac{w_{hs}}{w_{ls}}\right)$	$log\left(\frac{w_{hs}}{w_{ls}}\right)$	$log\left(\frac{w_{hs}}{w_{ls}}\right)$
$log\,(y_c)$	-0.360***	-0.415***	-0.251***	-0.250***
	(0.093)	(0.097)	(0.072)	(0.072)
$1_{[c\in 2]} \cdot log\,(y_c)$	0.047***	0.054***	0.022*	0.025*
	(0.017)	(0.018)	(0.013)	(0.013)
$1_{[c\in 3 \mid c\in 4]} \cdot log\,(y_c)$	0.067***	0.080***	0.045**	0.050***
	(0.024)	(0.025)	(0.018)	(0.019)
Constant	3.605***	4.096***	2.644***	2.576***
	(0.724)	(0.755)	(0.562)	(0.565)
Observations	80	80	80	80
R^2	0.229	0.252	0.187	0.160
Adjusted R^2	0.198	0.222	0.155	0.127
Prob > F	0.000181	5.94e-05	0.00123	0.00395

* Standard errors in parentheses (*** p<0.01, ** p<0.05, * p<0.1).

Extreme Values.

In addition to the robustness checks discussed above, I here study if my results are driven by the presence of extreme values in the occupational skill-premium. To that end, I re-estimate Model (3) in Table 1.2 taking out of my estimation sample the following countries, one at a time: Norway, Tanzania, Rwanda, Hong Kong, Laos, South Africa, Gambia, Tajikistan. I also explore how the results change if I drop all these countries together from my sample. In all cases, the results are robust to the exclusion of these countries. The estimated elasticities are both jointly and statistically significant at the one percent level [4]. The goodness of fit, as measured by the regression's R^2 improves for all cases, with the exception of the estimations that leave Norway and Laos out.

The Role of the Economic Environment and Institutions.

It is still an open discussion in the economic literature to what extent skill-premia is determined by worker's skills or attributes and how much of it depends on variables that affect the economic environment of countries, like the quality of institutions, openness to trade, their economic structure, and other cultural, organizational, or social norms in place [5].

A major concern related to this discussion is that the process of economic development is often characterized by significant improvements in the economic environment of countries. These improvements are sometimes driven by enhanced institutions, higher openness to trade, or other organizational. If those changes are neutral, in the sense that they do not affect workers of unlike skill types differently, they should not have an impact on the skill-

[4]The only exception is when I exclude Laos, where the coefficient for the second quartile is significant at the five percent level.

[5]For example, in Caselli and Ciccone (2019)'s words: *"it seems extremely implausible that attributes of workers are the sole determinant of skill-premia not accounted for by skill supply. Instead, it seems very likely that skill-premia are also shaped by institutions, technology, organizational structures, infrastructure, the structural composition of the economy, openness to trade, social norms, and other attributes of the environment."*

premium. On the other hand, if these types of changes affect workers of different skill types in heterogeneous ways, one should be cautious before ignoring their effects on skill-premium evolution across the development spectrum.

To attend these concerns, I study if the quantitative results presented in Subsection 1.3.1 remain robust after controlling for different sets of institutional, organizational, economic policy quality, and economic structure variables that might have an effect on the occupational skill-premium behavior.

To that purpose, I explore specifications that control for different set of institutional quality variables used in the literature. To be precise, I study the effects of three groups of institutional quality controls: the components of the Index Economic Freedom, the set of variables in the Worldwide Governance Indicators, and the institutional controls used by Acemoglu et al. (2014). The estimation results are presented in Table A.4 below, under labels Model (2.1)-(2.3).

Table A.5: Development Elasticity of the Occupational Skill-Premium
(robustness of non-linearities under institutional controls)

Variables	**Model (2.1)** $log\left(\frac{w_{hs}}{w_{ls}}\right)$	**Model (2.2)** $log\left(\frac{w_{hs}}{w_{ls}}\right)$	**Model (2.3)** $log\left(\frac{w_{hs}}{w_{ls}}\right)$	**Model (2.4)** $log\left(\frac{w_{hs}}{w_{ls}}\right)$
$log\,(y_c)$	-0.430***	-0.454***	-0.373***	-0.461***
	(0.128)	(0.107)	(0.118)	(0.099)
$1_{[c\in 2]} \cdot log\,(y_c)$	0.059***	0.050***	0.055***	0.043**
	(0.020)	(0.018)	(0.018)	(0.017)
$1_{[c\in 3 \mid c\in 4]} \cdot log\,(y_c)$	0.076***	0.070***	0.072***	0.064***
	(0.027)	(0.024)	(0.024)	(0.023)
X_{EFI}	[1]	-	-	-
	-	-	-	-
X_{WGI}	-	[2]	-	-
	-	-	-	-
Rule of Law	-	-	0.057	-
	-	-	(0.063)	-
Years of School.	-	-	-0.024	-
	-	-	(0.022)	-
Ind. VA Share	-	-	-	0.011
	-	-	-	(0.007)
Serv. VA Share	-	-	-	0.016**
	-	-	-	(0.006)
Constant	4.174***	4.424***	3.886***	3.418***
	(1.035)	(0.870)	(0.912)	(0.722)
Observations	79	80	77	80
R-squared	0.399	0.326	0.258	0.290
Adjusted R-squared	0.256	0.239	0.206	0.242
Prob > F	0.00229	0.000670	0.000621	9.55e-05

* Standard errors in parentheses (*** p<0.01, ** p<0.05, * p<0.1).

[1] X_{EFI} is a vector of controls composed by the sub-indeces in the Index of Economic Freedom Index, including: Property Rights, Government Integrity, Tax Burden, Government Spending, Fiscal Health, Business Freedom, Labor Freedom, Trade Freedom, Investment Freedom, and Financial Freedom. The Government Freedom component is statistically significant at the 1% level. All other components are not statistically significant at the 10% level.

[2] X_{WGI} is a vector of controls composed by the variables in the World Governance Indicators, including: Government Effectiveness, Political Stability and Absence of Violence, Regulatory Quality, Rule of Law,Voice and Accountability, and Control of Corruption. The Political Stability and Absence of Violence component is statistically significant at the 5% level.All other components are not statistically significant at the 10% level.

The Table shows that the results are robust after controlling for the three sets of institutional variables under consideration. The elasticity coefficients are all individually statistically significant at the one percent level and the Adjusted R^2 improves in the three cases, being Model (1) the best specification under this criterion. Quantitatively, the biggest change in the estimated elasticities compared to Model 3 in Table 1.2 is in Model 2.2 (-0.454,-0.404,-0.384), followed by Model 2.1 (-0.430,-0.371,-0.354), and Model 2.3 (-0.373,-0.318,-0.302).

Model (2.4) shows how the results change after controlling for variables that capture the economic structure of countries, namely, the share of Total Value Added in Industry and the share of Total Value Added in Services [6]. The results are, again, similar to the ones in Model (3), with the individual coefficients being all statistically significant at the one percent level. Quantitatively, the elasticities show the biggest change for all the Models presented in Table increasing in absolute value to 0.461,0.418, and 0.397, respectively [7].

Relative Total Labor Income versus Relative Hourly Labor Income.

I here study if my results are driven by different trends in hours worked across development between my major occupational groups. To that end, instead of studying the behavior of relative hourly labor income by broad occupational groups I focus on the behavior of relative total labor income in high- and low-skill occupations. Table A.6 below presents the same regressions as those in Table 1.2 in Subsection 1.3.1.

[6] I explored with different combinations of the share of Value Added in Agriculture, Industry, and Services. The elasticities are statistically significant at the one percent level in all cases. Model (2.4) presents the best specification I found.

[7] The results are also robust after jointly controlling for all the variables in the three institutional control groups and the economic structure variables.

Table A.6: Development Elasticity of the Occupational Skill-Premium
(relative occupational total labor income)

Variables	**Model (1)** $log\left(\frac{w_{hs}}{w_{ls}}\right)$	**Model (2)** $log\left(\frac{w_{hs}}{w_{ls}}\right)$	**Model (3)** $log\left(\frac{w_{hs}}{w_{ls}}\right)$
$log\,(y_c)$	-0.102**	-0.317*	-0.371***
	(0.046)	(0.162)	(0.132)
$1_{[c\in 2]} \cdot log\,(y_c)$	-	0.040	0.047 *
	-	(0.027)	(0.024)
$1_{[c\in 3]} \cdot log\,(y_c)$	-	0.066*	-
	-	(0.036)	-
$1_{[c\in 4]} \cdot log\,(y_c)$	-	0.057	-
	-	(0.045)	-
$1_{[c\in 3,4]} \cdot log\,(y_c)$	-	-	0.074**
	-	-	(0.034)
Constant	1.631***	3.219**	3.643***
	(0.435)	(1.262)	(1.033)
Observations	80	80	80
R-squared	0.059	0.119	0.115
Adjusted R-squared	0.047	0.072	0.083
Prob >F	0.0294	0.0467	0.0249

[*] Standard errors in parentheses (*** p<0.01, ** p<0.05, * p<0.1).

The main message from Table A.6 is that, qualitatively, the results are similar to the ones obtained in the relative hourly labor income. The best specification is still given by the model where elasticities vary with development until countries reach the third quartile, being all the estimated coefficients statistically significant at least at the ten percent level and jointly significant at the five percent level. Quantitatively, both the initial predicted premium and the estimated elasticities are, in absolute value, smaller. To be precise, the constant falls from 3.63 to 3.15, while the estimated elasticities are -0.315, -0.276, -0.253, compared to -0.362, -0.314, -0.295 in the case I use relative hourly labor income as my skill-premium measure.

The intuition behind this result becomes more clear after looking at the evolution of average weekly hours worked by major occupational groups across development, which I summarize in Table A.3 in Appendix A.3 [8]. As we can see, hours are initially higher in low-skill occupations and exhibit a higher decline as we move from the bottom to the top development quartile. These two effects together amplify the decline in the skill-premium for the following reasons: first, the fact that hours worked are higher in Low- than in high-skill occupations at low development levels drives the initial skill-premium up for the relative hourly labor income measure; second, the fact that hours worked decline at a faster pace in low-skill occupations as countries develop increases (in absolute terms) the GDP per capita elasticities the relative hourly labor income measure.

Workers of Both Sexes vs Males Only.

I here analyze to what extent my results depend on the inclusion or not of women when computing relative hourly labor income. I perform this robustness check since, for example, it might be a concern that women's attachment to the labor force or the gender wage gap vary with development.

I thus repeat my empirical analysis but now taking into account employment, hours worked and labor income data for male workers only. The regression results are presented in Table A.7 below.

As we can see, the results are in line with those obtained when considering workers of both sexes. The best specification is still given by the model where the elasticity varies with development and stabilizes in the third quartile, with the estimated coefficients being all significant at the one percent level.

[8] I also show in Appendix A.3 how ILO data on hours worked compares to others in the literature.

Table A.7: Development Elasticity of the Occupational skill-premium
(including males workers only)

Variables	**Model (1)** $log\left(\frac{w_{hs}}{w_{ls}}\right)$	**Model (2)** $log\left(\frac{w_{hs}}{w_{ls}}\right)$	**Model (3)** $log\left(\frac{w_{hs}}{w_{ls}}\right)$
$log\,(y_c)$	-0.061*	-0.300**	-0.321***
	(0.033)	(0.117)	(0.097)
$1_{[c\in 2]} \cdot log\,(y_c)$	-	0.046**	0.049***
	-	(0.020)	(0.018)
$1_{[c\in 3]} \cdot log\,(y_c)$	-	0.070**	-
	-	(0.027)	-
$1_{[c\in 4]} \cdot log\,(y_c)$	-	0.066**	-
	-	(0.033)	-
$1_{[c\in 3,4]} \cdot log\,(y_c)$	-	-	0.073***
	-	-	(0.025)
Constant	1.296***	3.076***	3.243***
	(0.314)	(0.922)	(0.757)
Observations	79	79	79
R-squared	0.042	0.143	0.142
Adjusted R-squared	0.030	0.097	0.107
Prob >F	0.0005	0.021	0.009

* Standard errors in parentheses (*** p<0.01, ** p<0.05, * p<0.1).

A.5 Proofs

A.5.1 Proposition 1

Under independent marginals and common shape parameter the share of workers in high-skill occupations is given by

$$
\begin{aligned}
\pi_h &= P\left(w_h Z_h \geq w_l Z_l\right) = P\left(Z_h \geq \frac{w_l}{w_h} Z_l\right) \\
&= \int_0^{+\infty} \int_0^{\left(\frac{w_h}{w_l}\right) z_h} g_l\left(z_l\right) g_h\left(z_h\right) dz_l dz_h \\
&= \int_0^{+\infty} e^{-S_l\left(z_h\left(\frac{w_h}{w_l}\right)\right)} g_h\left(z_h\right) dz_h \\
&= \int_0^{+\infty} e^{-S_l\left(z_h\left(\frac{w_h}{w_l}\right)\right)^{-\theta}} S_h z_h^{-\theta-1} e^{-S_h\left(z_h\right)^{-\theta}} dz_h \\
&= S_h \int_0^{+\infty} \theta z_h^{-\theta-1} e^{-\left(S_l\left(\frac{w_h}{w_l}\right)^{-\theta}+S_h\right) z_h^{-\theta}} dz_h \\
&= \frac{S_h}{\left(S_l\left(\frac{w_h}{w_l}\right)^{-\theta}+S_h\right)} \int_0^{+\infty} \theta\left(S_l\left(\frac{w_h}{w_l}\right)^{-\theta}+S_h\right) z_h^{-\theta-1} e^{-\left(S_l\left(\frac{w_h}{w_l}\right)^{-\theta}+S_h\right) z_h^{-\theta}} dz_h \\
&= \frac{S_h}{\left(S_l\left(\frac{w_h}{w_l}\right)^{-\theta}+S_h\right)}\left(\left. e^{-\left(S_l\left(\frac{w_h}{w_l}\right)^{-\theta}+S_h\right) z_h^{-\theta}} \right|_0^{+\infty}\right. = \frac{S_h}{\left(S_l\left(\frac{w_h}{w_l}\right)^{-\theta}+S_h\right)},
\end{aligned}
$$

and the distribution of labor productivity conditional on workers choosing high-skill occupations is

$$
\begin{aligned}
M_h\left(z\right) &= G_h\left(Z_h \leq z \mid w_h Z_h > w_l Z_l\right) \\
&= \int_0^z \int_0^{z_h\left(\frac{w_h}{w_l}\right)} g_h(z_h) g_l(z_l) dz_l dz_h \\
&= \frac{S_h}{\left(S_h + S_l\left(\frac{w_h}{w_l}\right)^{-\theta}\right)} \int_0^z \theta z_h^{-\theta-1} \left(S_h + S_l\left(\frac{w_h}{w_l}\right)^{-\theta}\right) e^{-\left(S_h + S_l\left(\frac{w_h}{w_l}\right)^{-\theta}\right) z_h^{-\theta}} dz_h \\
&= \pi_h \cdot e^{-\left(S_h + S_l\left(\frac{w_h}{w_l}\right)^{-\theta}\right) z^{-\theta}}.
\end{aligned}
$$

Average labor productivity in high-skill occupations is

$$
\begin{aligned}
\mathbb{E}\big(Z_h | w_h z_h > w_l z_l\big) &= \int_0^{+\infty} z_h M(z_h | w_h z_h > w_l z_l) dz_h, \\
&= \int_0^{+\infty} z_h \frac{g(z_h, w_h z_h > w_l z_l)}{g(w_h z_h > w_l z_l)} dz_h, \\
&= \left(\frac{1}{\pi_h}\right) \int_0^{+\infty} \int_0^{\left(\frac{w_h}{w_l}\right) z_h} z_h g(z_h, z_l) dz_l dz_h,
\end{aligned}
$$

in the independent case

$$
\mathbb{E}\big(Z_h | w_h z_h > w_l z_l\big) = \left(\frac{1}{\pi_h}\right) \int_0^{+\infty} \int_0^{\left(\frac{w_h}{w_l}\right) z_h} z_h g_h(z_h) g_l(z_l) dz_l dz_h.
$$

Under common shape parameter

$$
\begin{aligned}
\mathbb{E}\big(Z_h|w_hz_h > w_lz_l\big) &= \left(\frac{1}{\pi_h}\right)\int_0^{+\infty} z_h g_h(z_h) e^{-S_l\left(\frac{w_h}{w_l}z_h\right)^{-\theta}} dz_l dz_h,\\
&= \left(\frac{1}{\pi_h}\right)\int_0^{+\infty} z_h \theta S_h z_h^{(\theta-1)} e^{-S_h(z_h)^{-\theta}} e^{-S_l\left(\frac{w_h}{w_l}z_h\right)^{-\theta}} dz_h,\\
&= \int_0^{+\infty} z_h \theta z_h^{(\theta-1)} \left(S_h + S_l\left(\frac{w_h}{w_l}\right)^{-\theta}\right) e^{-\left(S_h+S_l\left(\frac{w_h}{w_l}\right)^{-\theta}\right)z_h^{-\theta}} dz_h,\\
&= \int_0^{+\infty} z_h \theta z_h^{(\theta-1)} \left(S_h + S_l\left(\frac{w_h}{w_l}\right)^{-\theta}\right) e^{-\left(S_h+S_l\left(\frac{w_h}{w_l}\right)^{-\theta}\right)z_h^{-\theta}} dz_h,
\end{aligned}
$$

Let $y = \left(S_h + S_l\left(\frac{w_h}{w_l}\right)^{-\theta}\right) z_h^{-\theta}$. It follows that

$$
\begin{aligned}
\mathbb{E}\big(Z_h|w_hz_h > w_lz_l\big) &= \int_0^{+\infty} \left(\frac{y}{S_h + S_l\left(\frac{w_h}{w_l}\right)^{-\theta}}\right)^{-\left(\frac{1}{\theta}\right)} e^{-y} dy,\\
\mathbb{E}\big(Z_h|w_hz_h > w_lz_l\big) &= \left(S_h + S_l\left(\frac{w_h}{w_l}\right)^{-\theta}\right)^{\left(\frac{1}{\theta}\right)} \Gamma\left(1-\frac{1}{\theta}\right) \qquad units
\end{aligned}
$$

$\square$

A.5.2 Proposition 2

(Educational Sorting Conditional on Occupational Choice). Let c^s and c^u denote the fixed costs of acquiring Secondary and University education, β^s the log-return of completing Secondary education and β_h^u and β_l^u the log-returns to completing University education in high- and low-skill occupations, respectively. Assume $c^u > c^s > 0$ and $\beta_h^u > \beta_l^u > \beta^s > 0$.
Conditional on working in high-skill occupations, workers' incomes after educational costs are:

1. **High-skill occupations, No-Schooling:** $W_h^{ns} = w_h z_h$

2. **High-skill occupations, Secondary Complete:** $W_h^s - c^s = w_h z_h e^{\beta^s} - c^s$

3. **High-skill occupations, University Complete:** $W_h^u - c^u = w_h z_h e^{\beta_h^u} - c^u$

Workers choose no-schooling if:

$$
\begin{aligned}
W_h^{ns} &\geq W_h^s - c^s \\
w_h z_h &\geq w_h z_h e^{\beta^s} - c^s \\
z_h &\leq \frac{c^s}{w_h \left(e^{\beta^s} - 1\right)} = z_h^{\star},
\end{aligned}
$$

and

$$
\begin{aligned}
W_h^{ns} &\geq W_h^u - c^u \\
w_h z_h &\geq w_h z_h e^{\beta_h^u} - c^u \\
z_h &\leq \frac{c^u}{w_h \left(e^{\beta_h^u} - 1\right)} = \tilde{z}_h^{\star}.
\end{aligned}
$$

Secondary Education is chosen if

$$
\begin{aligned}
W_h^s - c^s &> W_h^{ns} \\
w_h z_h e^{\beta^s} - c^s &> w_h z_h \\
z_h &> \frac{c^s}{w_h\left(e^{\beta^s} - 1\right)} = z_h^\star,
\end{aligned}
$$

and

$$
\begin{aligned}
W_h^s - c^s &\geq W_h^u - c^u \\
w_h z_h e^{\beta^s} - c^s &\geq w_h z_h e^{\beta_h^u} - c^u \\
z_h &\leq \frac{c^u - c^s}{w_h\left(e^{\beta_h^u} - e^{\beta^s}\right)} = z_h^{\star\star}.
\end{aligned}
$$

Finally, still conditional on working in high-skill occupations, individuals choose to acquire University schooling if

$$
\begin{aligned}
W_h^u - c^u &> W_h^{ns} \\
w_h z_h e^{\beta_h^u} - c^u &> w_h z_h \\
z_h &> \frac{c^u}{w_h\left(e^{\beta_h^u} - 1\right)} = \tilde{z}_h^\star,
\end{aligned}
$$

and

$$
\begin{aligned}
W_h^u - c^u &> W_h^s - c^s \\
w_h z_h e^{\beta_h^u} - c^u &> w_h z_h e^{\beta^s} - c^s \\
z_h &> \frac{c^u - c^s}{w_h\left(e^{\beta_h^u} - e^{\beta^s}\right)} = z_h^{\star\star}
\end{aligned}
$$

To guarantee that the educational attainment rule is monotonically increasing in worker's

ability, one needs $z_h^\star < \tilde{z}_h^\star < z_h^{\star\star}$. Otherwise, if $\tilde{z}_h^\star < z_h^\star$, workers with ability $z_h \in [0, \tilde{z}_h^\star]$ choose no schooling over Secondary school, workers with ability $z_h \in (\tilde{z}_h^\star, z_h^\star]$ choose University over Secondary, individuals with $z_h \in (z_h^\star, z_h^{\star\star}]$ choose Secondary over University and no schooling, and workers with $z_h > z_h^{\star\star}$ choose University education. Following the same logic, if $\tilde{z}_h^\star > z_h^{\star\star}$, the educational attainment rule is monotonically increasing in ability until $z_h^{\star\star}$, where a region of workers with $z_h \in (z_h^{\star\star}, \tilde{z}_h^\star)$ emerges and where workers prefer Secondary Schooling to no schooling, University schooling to Secondary schooling and no schooling over University schooling could emerge for workers with abilities $z_h \in (z_h^{\star\star}, \tilde{z}_h^\star)$.
After some algebra, a sufficient condition for $z_h^\star < \tilde{z}_h^\star < z_h^{\star\star}$ to hold is

$$\left(\frac{c^u}{c^s}\right) > \left(\frac{\beta_h^u - 1}{\beta^s - 1}\right)$$

Conditional on working in low-skill occupations, workers' incomes after educational costs are:

1. **No-Schooling:** $W_l^{ns} = w_l z_l$
2. **Secondary Complete:** $W_l^s - c^s = w_l z_l e^{\beta^s} - c^s$
3. **University Complete:** $W_l^u - c^u = w_l z_l e^{\beta_l^u} - c^u$

Following the same logic, the ability thresholds to choose Secondary education over no schooling $z_l^\star$ and University over Secondary education $z_l^{\star\star}$ conditional on working in low-skill occupations are given by

$$z_l^\star = \frac{c^s}{w_l\left(e^{\beta^s} - 1\right)}$$
$$z_l^{\star\star} = \frac{c^u - c^s}{w_l\left(e^{\beta_l^u} - e^{\beta^s}\right)}.$$

A sufficient condition for the educational decision rule to be monotonically increasing in workers' ability in low-skill occupations is

$$\left(\frac{c^u}{c^s}\right) > \left(\frac{\beta_l^u - 1}{\beta^s - 1}\right),$$

which, given that $\beta_h^u > \beta_l^u$ is guaranteed by $\left(\frac{c^u}{c^s}\right) > \left(\frac{\beta_h^u - 1}{\beta^s - 1}\right)$.

$\square$

A.5.3 Proposition 3

Define $\tilde{z}_l^{\star\star} = z_h^{\star\star}\left(\frac{w_h}{w_l}\right) = \frac{c^u - c^s}{w_l\left(e^{\beta_h^u} - e^{\beta^s}\right)}$. Under endogenous human capital accumulation through schooling, the occupational selection decisions are:

1. If $z_l \leq z_l^\star$ workers choose high-skill occupations if $z_h > \left(\frac{w_l}{w_h}\right) z_l$.

2. If $z_l \in (z_l^\star, \tilde{z}_l^{\star\star}]$ workers choose high-skill occupations if $z_h > \left(\frac{w_l}{w_h}\right) z_l$.

3. If $z_l \in (\tilde{z}_l^{\star\star}, z_l^{\star\star}]$ workers choose high-skill occupations if $z_h > \left(\frac{w_l}{w_h}\right)\left(\frac{z_l e^{\beta^s}}{e^{\beta_h^u}}\right)$.

4. If $z_l > z_l^{\star\star}$ workers choose high-skill occupations if $z_h > \left(\frac{w_l}{w_h}\right)\left(\frac{z_l e^{\beta_l^u}}{e^{\beta_h^u}}\right)$.

The proof goes by contraposition.

1. Suppose not. Then, there exists a $z_l \in [0, z_l^\star]$ such that $w_h z_h \leq w_l z_l$ and workers choose to work in high-skill occupations.

Then there exists a $z_l \in [0, z_l^\star]$ such that

$$w_h z_h \leq w_l z_l$$

and

$$\max_{\{W_h^{ns}, W_h^s, W_h^u\}} \left\{W_h^{ns}, W_h^s - c^s, W_h^u - c^u\right\} > \max_{\{W_l^{ns}, W_l^s, W_l^u\}} \left\{W_l^{ns}, W_l^s - c^s, W_l^u - c^u\right\}.$$

Since $z_l \in [0, z_l^\star]$, $\max\limits_{\{W_l^{ns}, W_l^s, W_l^u\}} \left\{ W_l^{ns}, W_l^s - c^s, W_l^u - c^u \right\} = W_l^{ns} = w_l z_l$.

If $\max\limits_{\{W_h^{ns}, W_h^s, W_h^u\}} \left\{ W_h^{ns}, W_h^s - c^s, W_h^u - c^u \right\} = W_h^{ns} \implies w_h z_h > w_l z_l$, which is a contradiction.

If $\max\limits_{\{W_h^{ns}, W_h^s, W_h^u\}} \left\{ W_h^{ns}, W_h^s - c^s, W_h^u - c^u \right\} = W_h^s = w_h z_h e^{\beta^s}$, $z_h \in (z_h^\star, z_h^{\star\star}]$, which together with $w_l z_l \geq w_h z_h \implies z_l > \left(\frac{w_h}{w_l}\right) z_h^\star = z_l^\star$, which is a contradiction.

If $\max\limits_{\{W_h^{ns}, W_h^s, W_h^u\}} \left\{ W_h^{ns}, W_h^s - c^s, W_h^u - c^u \right\} = W_h^u = w_h z_h e^{\beta_h^u}$, $z_h \in (z_h^{\star\star}, +\infty]$, which together with $w_l z_l \geq w_h z_h \implies z_l > \left(\frac{w_h}{w_l}\right) z_h^{\star\star} = \tilde{z}_l^{\star\star} > z_l^\star$, which is a contradiction.

2. Suppose not. Then, there exists a $z_l \in (z_l^\star, \tilde{z}_l^{\star\star}]$ such that $w_h z_h \leq w_l z_l$ and workers choose to work in high-skill occupations.

Then there exists a $z_l \in (z_l^\star, \tilde{z}_l^{\star\star}]$ such that

$$w_h z_h \leq w_l z_l$$

and

$$\max_{\{W_h^{ns}, W_h^s, W_h^u\}} \left\{ W_h^{ns}, W_h^s - c^s, W_h^u - c^u \right\} > \max_{\{W_l^{ns}, W_l^s, W_l^u\}} \left\{ W_l^{ns}, W_l^s - c^s, W_l^u - c^u \right\}.$$

Since $z_l \in (z_l^\star, \tilde{z}_l^{\star\star}]$, $\max\limits_{\{W_l^{ns}, W_l^s, W_l^u\}} \left\{ W_l^{ns}, W_l^s - c^s, W_l^u - c^u \right\} = W_l^s = w_l z_l e^{\beta^s}$.

If $\max\limits_{\{W_h^{ns}, W_h^s, W_h^u\}} \left\{ W_h^{ns}, W_h^s - c^s, W_h^u - c^u \right\} = W_h^{ns} \implies w_h z_h > w_l z_l e^{\beta^s} > w_l z_l$, which is a contradiction.

If $\max\limits_{\{W_h^{ns}, W_h^s, W_h^u\}} \left\{ W_h^{ns}, W_h^s - c^s, W_h^u - c^u \right\} = W_h^s = w_h z_h e^{\beta^s} \implies w_h z_h e^{\beta^s} > w_l z_l e^{\beta^s} \implies w_h z_h > w_l z_l$, which is a contradiction.

3. Suppose not. Then, there exists a $z_l \in (\tilde{z}_l^{\star\star}, z_l^{\star\star}]$ such that $w_h z_h e^{\beta_h^u} \leq w_l z_l e^{\beta^s}$ and workers choose to work in high-skill occupations.

Then there exists a $z_l \in (\tilde{z}_l^{\star\star}, z_l^{\star\star}]$ such that

$$w_h z_h e^{\beta_h^u} \leq w_l z_l e^{\beta^s}$$

and

$$\max_{\{W_h^{ns}, W_h^s, W_h^u\}} \left\{W_h^{ns}, W_h^s - c^s, W_h^u - c^u\right\} > \max_{\{W_l^{ns}, W_l^s, W_l^u\}} \left\{W_l^{ns}, W_l^s - c^s, W_l^u - c^u\right\}.$$

Since $z_l \in (\tilde{z}_l^{\star\star}, z_l^{\star\star}]$, $\max_{\{W_l^{ns}, W_l^s, W_l^u\}} \left\{W_l^{ns}, W_l^s - c^s, W_l^u - c^u\right\} = W_l^s = w_l z_l e^{\beta^s}$. Also, $w_h z_h e^{\beta_h^u} \leq w_l z_l e^{\beta^s} \implies w_h z_h \leq w_l z_l$.

If $\max_{\{W_h^{ns}, W_h^s, W_h^u\}} \left\{W_h^{ns}, W_h^s - c^s, W_h^u - c^u\right\} = W_h^{ns}$,

$$w_h z_h > w_l z_l e^{\beta^s}$$

$$w_h z_h > w_l z_l, \text{ which is a contradiction.}$$

If $\max_{\{W_h^{ns}, W_h^s, W_h^u\}} \left\{W_h^{ns}, W_h^s - c^s, W_h^u - c^u\right\} = W_h^s$,

$$w_h z_h e^{\beta^s} > w_l z_l e^{\beta^s}$$

$$w_h z_h > w_l z_l, \text{ which is a contradiction.}$$

If $\max_{\{W_h^{ns}, W_h^s, W_h^u\}} \left\{W_h^{ns}, W_h^s - c^s, W_h^u - c^u\right\} = W_h^u$,

$$w_h z_h e^{\beta_h^u} > w_l z_l e^{\beta^s} \text{ which is a contradiction.}$$

4. Suppose not. Then, there exists a $z_l \in (z_l^{\star\star}, +\infty)$ such that $w_h z_h e^{\beta_h^u} \leq w_l z_l e^{\beta_l^u}$ and

workers choose to work in high-skill occupations.

Then there exists a $z_l \in (z_l^\star, \tilde{z}_l^{\star\star})$ such that

$$w_h z_h e^{\beta_h^u} \leq w_l z_l e^{\beta_l^u}$$

and

$$\max_{\{W_h^{ns}, W_h^s, W_h^u\}} \left\{W_h^{ns}, W_h^s - c^s, W_h^u - c^u\right\} > \max_{\{W_l^{ns}, W_l^s, W_l^u\}} \left\{W_l^{ns}, W_l^s - c^s, W_l^u - c^u\right\}.$$

Since $z_l \in (z_l^{\star\star}, +\infty)$, $\max_{\{W_l^{ns}, W_l^s, W_l^u\}} \left\{W_l^{ns}, W_l^s - c^s, W_l^u - c^u\right\} = W_l^u = w_l z_l e^{\beta_l^u}$. Also, $w_h z_h e^{\beta_h^u} \leq w_l z_l e^{\beta_h^l} \implies w_h z_h \leq w_l z_l$.

If $\max_{\{W_h^{ns}, W_h^s, W_h^u\}} \left\{W_h^{ns}, W_h^s - c^s, W_h^u - c^u\right\} = W_h^{ns}$,

$$w_h z_h > w_l z_l e^{\beta_l^u}$$

$$w_h z_h > w_l z_l, \text{ which is a contradiction.}$$

If $\max_{\{W_h^{ns}, W_h^s, W_h^u\}} \left\{W_h^{ns}, W_h^s - c^s, W_h^u - c^u\right\} = W_h^s$,

$$w_h z_h e^{\beta^s} > w_l z_l e^{\beta_l^u}$$

$$w_h z_h > w_l z_l, \text{ which is a contradiction.}$$

If $\max_{\{W_h^{ns}, W_h^s, W_h^u\}} \left\{W_h^{ns}, W_h^s - c^s, W_h^u - c^u\right\} = W_h^u$,

$$w_h z_h e^{\beta_h^u} > w_l z_l e^{\beta^s} \text{ which is a contradiction.}$$

A.5.4 Proposition 4

The share of workers in high-skill occupations is given by:

$$\pi_h = \int_0^{z_h^{\star\star}} \int_0^{\left(\frac{w_h}{w_l}\right) z_h} g_\phi\left(z_h, z_l\right) dz_h dz_l + \int_{z_h^{\star\star}}^{\tilde{z}_h^{\star\star}} \int_0^{\left(\frac{w_h}{w_l}\right)\left(\frac{e^{\beta_h^u}}{e^{\beta^s}}\right) z_h} g_\phi\left(z_h, z_l\right) dz_h dz_l$$
$$+ \int_{\tilde{z}_h^{\star\star}}^{+\infty} \int_0^{\left(\frac{w_h}{w_l}\right)\left(\frac{e^{\beta_h^u}}{e^{\beta_l^u}}\right) z_h} g_\phi\left(z_h, z_l\right) dz_h dz_l$$

The proof follows from Proposition 3. The first term on the right hand side captures all the workers who choose high-skill occupations if $z_h > \left(\frac{w_l}{w_h}\right) z_l$. The limits of the outer integral capture the abilities in high-skill occupations that make workers indifferent between choosing high- and low-skill occupations in the limits of the interval, namely $[0, z_h^{\star\star}]$ (and $[0, \tilde{z}_l^{\star\star}]$) . The limits of the inner integral captures all the individuals who have ability below the indifference level in low-skill occupations for a given ability level in high-skill occupations, namely, the $z_l \in \left[0, \left(\frac{w_h}{w_l}\right) z_h\right]$.

The second and third terms on the right hand side of the equation are analogous, with the only difference that the second term captures the ability region were workers get Secondary education in low-skill occupations and University education in high-skill occupations, and the third term captures the ability region were workers get University education in high- and low-skill occupations.

A.6 Census/American Community Survey data issues and handling.

To discipline the innate ability distribution I use IPUMS International data for the US in 2010, which contains Household Survey micro-data from the American Community survey, harmonized to allow for international comparisons. The sample includes 1% of the United State's population.

In order to minimize noise in the calculations, wages are computed only for workers with a considerable attachment to the labor force, following Acemoglu and Autor (2011a) criteria. Thus, I consider only full-time (i.e. at least 35 hours per week), full year (40 weeks per year or more) workers, aged 16-64, and exclude those who are in the military, institutionalized, or self employed. I construct hourly earnings as the ratio of annual earnings and total hours worked, being the latter the product of average weeks worked per year and average hours worked per week. Calculations are weighted by ACS sampling weights and are converted in real terms using the personal consumer expenditure (PCE) deflator. Eearnings below US$ 1.675 per hour (US$ 67 per week in Acemoglu and Author over 40 hours per week) in 1982 dollars are dropped. I replace income for top-coded earners with 1.45 times the value assigned to the corresponding top-level income, which in 2010 requires identifying the 99.5th percentile of income by state.

Separating occupational labor income from occupational wages that are common across workers and occupations requires and identifying assumption. To that end, I first classify occupations into two broad groups, high-, and low-skill, by following the procedure described in Section 1.2. I then assume that the average labor productivity of workers with no experience and no educational attainment in high- and low-skill occupations equals unity. Thus, the average labor income of the individuals in these two groups give me efficiency wages in high- and low-skill occupations. Labor productivity for the individuals in high- and low-skill occupations that do not belong to the base groups is obtained by dividing their labor income by the corresponding occupational efficiency wages.

Appendix B

Supplementary Material to Chapter 3

B.1 Proofs

B.1.1 Proposition 1

The Bellman Equation in the *normal phase* is given by

$$\rho V_N \frac{w^{1-\gamma}}{1-\gamma} = \max_{c,\alpha,\kappa,b} \left\{ c^{1-\gamma} \frac{w^{1-\gamma}}{1-\gamma} + V_N w^{1-\gamma} \left[(\alpha A + (1-\alpha) H) - (\kappa + b + c) \right] \right. \\ \left. -\gamma V_N w^{1-\gamma} \frac{\sigma^2}{2} \alpha^2 + \eta \left[V_D \frac{w^{1-\gamma}}{1-\gamma} E \left[\left(\beta(\alpha, b, \kappa, \mu^\delta, \delta, \zeta) \right)^{1-\gamma} \right] - V_N \frac{w^{1-\gamma}}{1-\gamma} \right] \right\}. \tag{B.1}$$

The FOCs are

$c_N : c_N^{-\gamma} - V_N w^{1-\gamma} = 0$

$$\alpha_N : (A - H) V_N w^{1-\gamma} - \sigma^2 \alpha_N \gamma V_N w^{1-\gamma}$$

$$+ \eta V_D w^{1-\gamma} \beta(\alpha, b, \kappa, \mu^\delta, \delta, \zeta)^{-\gamma} \left\{ \mu_k^\delta (\kappa) \delta_k (1 +_k) - \mu_h^\delta (\kappa) \delta_h (1 +_h) \right\} = 0$$

$$b_N : -V_N w^{1-\gamma} + \eta V_D w^{1-\gamma} \beta(\alpha, b, \kappa, \mu^\delta, \delta, \zeta)^{-\gamma} \left(\frac{\partial I(b)}{\partial b} \right) = 0$$

Re-arranging the FOCs we obtain

$$c_N = V_N^{-\left(\frac{1}{\gamma}\right)}, \tag{B.2}$$

and

$$\alpha_N = \frac{A - H}{\gamma \sigma^2} + \frac{\eta \Delta_k(\kappa)}{V \beta(\alpha_N, b, \kappa, \mu^\delta, \delta, \zeta)^\gamma}, \frac{1}{\gamma \sigma^2} \tag{B.3}$$

where we have defined $\Delta_k(\kappa)$ to be

$$\Delta_k(\kappa) = \mu_k^\delta(\kappa)(1 + \zeta_k) - \mu_h^\delta(\kappa)(1 + \zeta_h).$$

Under actuarially fairly priced insurance $I(b) = \frac{b}{\eta}$, $\frac{\partial I(b)}{\partial b} = \frac{1}{\eta}$, and optimal demand for disaster insurance implies

$$V \beta(\alpha_N, b, \kappa, \mu^\delta, \delta, \zeta)^\gamma = 1. \tag{B.4}$$

The Bellman Equation in the *disaster phase* is given by

$$\rho V_D \frac{w^{1-\gamma}}{1-\gamma} = \max_{c,\alpha,\kappa_D,b_D} \Bigg\{ c^{1-\gamma}\frac{w^{1-\gamma}}{1-\gamma} + V_D w^{1-\gamma}\left[\left(\alpha A_D + (1-\alpha) H_D\right) - (\kappa_D + c + b_D)\right]$$
$$- \gamma V_D w^{1-\gamma}\frac{\sigma^2}{2}\alpha^2 + \upsilon(\kappa_D)\left[V_N \frac{w^{1-\gamma}}{1-\gamma} - V_D \frac{w^{1-\gamma}}{1-\gamma}\right]$$
$$+ \eta\left[V_D \frac{w^{1-\gamma}}{1-\gamma} E\left[\left(\beta(\alpha, b_D, \kappa, \mu^\delta, \delta, \zeta_D)\right)^{1-\gamma}\right] - V_D \frac{w^{1-\gamma}}{1-\gamma}\right] \Bigg\} \tag{B.5}$$

The FOCs are

$c_D : c_D^{-\gamma} - V_D w^{1-\gamma} = 0$

$$\alpha_D : \left(A_D - H_D\right) V_D w^{1-\gamma} - \sigma^2 \alpha_D \gamma V_D w^{1-\gamma}$$
$$+ \eta V_D w^{1-\gamma} \beta(\alpha, b_D, \kappa, \mu^\delta, \delta, \zeta_D)^{-\gamma}\left\{\mu_k^\delta\left(\kappa\right)\delta_k\left(1+_k\right) - \mu_h^\delta\left(\kappa\right)\delta_h\left(1+_h\right)\right\} = 0$$

$$b_D : -V_D w^{1-\gamma} + \eta V_D w^{1-\gamma}\beta(\alpha, b_D, \kappa, \mu^\delta, \delta, \zeta_D)^{-\gamma}\left(\frac{\partial I\left(b\right)}{\partial b}\right) = 0$$

Re-arranging the FOCs we obtain

$$c_D = V_D^{-\left(\frac{1}{\gamma}\right)}, \tag{B.6}$$

and

$$\alpha_D = \frac{A_D - H_D}{\gamma\sigma^2} + \frac{\eta\Delta_k(\kappa)}{\beta(\alpha_D, b_D, \kappa, \mu^\delta, \delta, \zeta_D)^\gamma}\frac{1}{\gamma\sigma^2}. \tag{B.7}$$

Optimal demand for disaster insurance implies

$$\beta(\alpha_D, b_D, \kappa, \mu^\delta, \delta, \zeta_D) = 1. \tag{B.8}$$

Replacing (B.2) and (B.3) into (B.1), and dividing both sides by $V_N \frac{w^{1-\gamma}}{1-\gamma}$ we obtain

$$(\rho+\eta) = \Bigg\{\gamma V_N^{-\left(\frac{1}{\gamma}\right)} + (1-\gamma)\left[\alpha_N^\star \frac{(A-H)}{\gamma\sigma^2} + H - b_N^\star - \kappa_N^\star\right] - (1-\gamma)\gamma\frac{\sigma^2}{2}\left(\alpha_N^\star\right)^2$$
$$+ \eta\frac{V_D}{V_N}\left(\beta(\alpha_N^\star, b_N^\star, \kappa^\star, \mu^\delta, \delta, \zeta_N^\star)\right)^{1-\gamma}\Bigg\}$$

Similarly, replacing (B.6) and (B.7) into (B.5, and dividing both sides by $V_D \frac{w^{1-\gamma}}{1-\gamma}$ we obtain

$$(\rho+\eta+\nu(\kappa_D)) = \Bigg\{ \gamma V_D^{-\left(\frac{1}{\gamma}\right)} + (1-\gamma)\left[\alpha_D^\star \frac{(A_D - H_D)}{\gamma\sigma^2} + H_D - b_D^\star - \kappa_D^\star\right] - (1-\gamma)\gamma\frac{\sigma^2}{2}\left(\alpha_D^\star\right)^2$$

$$+\nu(\kappa_D)\frac{V_N}{V_D} + \eta\left(\beta(\alpha_D^\star, b_D^\star, \kappa^\star, \mu^\delta, \delta, \zeta_D^\star)\right)^{1-\gamma}\Bigg\}$$

To economize on notation we define

$$P = H + \alpha_N(A-H) - \gamma\frac{\sigma^2}{2}\alpha_N^2 - \kappa,$$

$$P_D = H_D + \alpha_D(A_D - H_D) - \gamma\frac{\sigma^2}{2}\alpha_D^2 - \kappa_D,$$

and

$$V = \frac{V_N}{V_D}.$$

V_N is pinned down by

$$\gamma V_N^{-\left(\frac{1}{\gamma}\right)} = (\rho+\eta) - (1-\gamma)P + (1-\gamma)b_N^\star - \eta V^{-1}\left(\beta(\alpha_N^\star, b_N^\star, \kappa^\star, \mu^\delta, \delta, \zeta_N^\star)\right)^{1-\gamma}, \quad \text{(B.9)}$$

and V_D by

$$\gamma V_D^{-\left(\frac{1}{\gamma}\right)} = (\rho+\eta+\nu(\kappa_D)) - (1-\gamma)P_D + (1-\gamma)b_D^\star - \eta\left(\beta(\alpha_D^\star, b_D^\star, \kappa^\star, \mu^\delta, \delta, \zeta_D^\star)\right)^{1-\gamma} - \nu(\kappa_D)V. \quad \text{(B.10)}$$

Dividing (B.10) by (B.9), we obtain

$$V^{\left(\frac{1}{\gamma}\right)} = \frac{(\rho+\eta+\nu(\kappa_D)) - (1-\gamma)P_D + (1-\gamma)b_D^\star - \eta\left(\beta(\alpha_D^\star, b_D^\star, \kappa^\star, \mu^\delta, \delta, \zeta_D^\star)\right)^{1-\gamma} - \nu(\kappa_D)V}{(\rho+\eta) - (1-\gamma)P + (1-\gamma)b_N^\star - \eta V^{-1}\left(\beta(\alpha_N^\star, b_N^\star, \kappa^\star, \mu^\delta, \delta, \zeta_N^\star)\right)^{1-\gamma}},$$

and the equilibrium relative valuation V^* is pinned down by

$$V^{\left(\frac{1}{\gamma}\right)}\left((\rho+\eta)-(1-\gamma)P+(1-\gamma)b_N^\star\right)-\eta V^{\frac{1-\gamma}{\gamma}}\left(\beta(\alpha_N^\star,b_N^\star,\kappa^\star,\mu^\delta,\delta,\zeta_N^\star)\right)^{1-\gamma}=$$
$$(\rho+\eta+\nu\left(\kappa_D\right))-(1-\gamma)P_D+(1-\gamma)b_D^\star-\eta\left(\beta(\alpha_D^\star,b_D^\star,\kappa^\star,\mu^\delta,\delta,\zeta_D^\star)\right)^{1-\gamma}-\nu\left(\kappa_D\right)V \tag{B.11}$$

Finally, the growth rates in the *normal* and *disaster* phases, conditional on no strikes, are:

$$\begin{aligned}\mu^N=\mathbb{E}\left(\frac{dw}{w}\mid N_t=0\right)&=\alpha_N\left(A-H\right)+H-\kappa+b_N+c_N\\&=H+\alpha_N\left(A-H\right)-\kappa-b_N-V_N^{-\left(\frac{1}{\gamma}\right)},\end{aligned}$$

and

$$\begin{aligned}\mu^D=\mathbb{E}\left(\frac{dw}{w}|N_t=0,M_t=0\right)&=\alpha_D\left(A_D-H_D\right)+H_D-\kappa_D+b_D+c_D\\&=H_D+\alpha_D\left(A_D-H_D\right)-\kappa_D-b_D-V_D^{-\left(\frac{1}{\gamma}\right)}.\end{aligned}$$

□

B.1.2 Existence, Uniqueness, and Properties of the Equilibrium.

Consider the case where the properties of natural disasters do not affect the composition of the portfolio. To be precise, assume

$$\mu^\delta(\kappa)=\mu_k^\delta(\kappa)(1+\zeta_k)=\mu_h^\delta(\kappa)(1+\zeta_h).$$

In this case, $\Delta_k\left(\kappa\right)=0$, and

$$\alpha_N=\frac{A-H}{\gamma\sigma^2},$$

$$\alpha_D = \frac{A_D - H_D}{\gamma\sigma^2},$$

and

$$\beta(\alpha, b, \kappa, \mu^\delta, \delta, \zeta) = \mu^\delta(\kappa) + \frac{b_N}{\eta} + \zeta_N.$$

Re-arranging (B.11), the relative valuation of the problem solves

$$\begin{aligned}
&(\rho + \eta) - (1-\gamma)P - \eta V^{-1}\beta(\alpha, b, \kappa, \mu^\delta, \delta, \zeta)^{1-\gamma} \\
&= \left[\rho + \eta + \upsilon(\kappa_D) - (1-\gamma)P_D - \eta\beta(\alpha, b, \kappa, \mu^\delta, \delta, \zeta_D)^{1-\gamma} - \upsilon(\kappa_D)V\right] V^{-1/\gamma}.
\end{aligned}$$

To study the properties of the solution, it is convenient to analyze the left and the right hand side of this equation separately. Define $L(V)$ and $R(V)$ to be

$$L(V) = (\rho + \eta) - (1-\gamma)P - \eta V^{-1}\beta(\alpha, b, \kappa, \mu^\delta, \delta, \zeta)^{1-\gamma},$$

and

$$R(V) = \left[\rho + \eta + \upsilon(\kappa_D) - (1-\gamma)P_D - \eta\beta(\alpha, b, \kappa, \mu^\delta, \delta, \zeta_D)^{1-\gamma} - \upsilon(\kappa_D)V\right] V^{-1/\gamma}.$$

$L(V)$ has the following properties:

- $L^{'}(V) > 0$
- $L^{''}(V) < 0$
- $\underset{V\to 0}{lim} L(V) \to -\infty$
- $\underset{V\to +\infty}{lim} L(V) \to (\rho + \eta) - (1-\gamma)P$
- $L(1) = (\rho + \eta) - (1-\gamma)P - \eta\beta(\alpha, b, \kappa, \mu^\delta, \delta, \zeta)^{1-\gamma}$

To study the properties of $R\left(V\right)$ it is convenient to express it as

$$R\left(V\right)=\left(\rho+\eta+\upsilon(\kappa_{D})-(1-\gamma)P_{D}-\eta\beta(\alpha,b,\kappa,\mu^{\delta},\delta,\zeta_{D})^{1-\gamma}\right)V^{-1/\gamma}-\upsilon(\kappa_{D})V^{\frac{\gamma-1}{\gamma}}.$$

If $\gamma>1$ and $\left(\rho+\eta+\upsilon(\kappa_{D})+(\gamma-1)P_{D}-\eta\beta(\alpha,b,\kappa,\mu^{\delta},\delta,\zeta_{D})^{1-\gamma}\right)>0$:

- $R^{'}\left(V\right)<0$
- $R^{''}\left(V\right)<0$
- $\underset{V\to0}{lim}R\left(V\right)\to+\infty$
- $\underset{V\to+\infty}{lim}R\left(V\right)\to-\infty$
- $R\left(1\right)=\left(\rho+\eta\right)+(\gamma-1)P_{D}-\eta\beta(\alpha,b,\kappa,\mu^{\delta},\delta,\zeta_{D})^{1-\gamma}.$

On the other hand, if $\gamma\in(0,1)$ and

$$\left(\rho+\eta-(1-\gamma)P_{D}-\eta\beta(\alpha,b,\kappa,\mu^{\delta},\delta,\zeta_{D})^{1-\gamma}\right)>0$$

- $R\left(V\right)>0\Leftrightarrow\left(\rho+\eta+\upsilon(\kappa_{D})-(1-\gamma)P_{D}-\eta\beta(\alpha,b,\kappa,\mu^{\delta},\delta,\zeta_{D})^{1-\gamma}\right)V^{-1/\gamma}>\upsilon(\kappa_{D})V^{1-\frac{1}{\gamma}}$, which requires:

$$V>\left(\frac{\rho+\eta+\upsilon(\kappa_{D})-(1-\gamma)P_{D}-\eta\beta(\alpha,b,\kappa,\mu^{\delta},\delta,\zeta_{D})^{1-\gamma}}{\upsilon(\kappa_{D})}\right)\equiv\tilde{V}.$$

- $R^{'}\left(V\right)>0\Leftrightarrow\upsilon(\kappa_{D})(1-\gamma)V^{-1}>\left(\rho+\eta+\upsilon(\kappa_{D})-(1-\gamma)P_{D}-\eta\beta(\alpha,b,\kappa,\mu^{\delta},\delta,\zeta_{D})^{1-\gamma}\right)$, which requires:

$$V>\left(\frac{\rho+\eta+\upsilon(\kappa_{D})-(1-\gamma)P_{D}-\eta\beta(\alpha,b,\kappa,\mu^{\delta},\delta,\zeta_{D})^{1-\gamma}}{\upsilon(\kappa_{D})(1-\gamma)}\right)\equiv\hat{V}$$

- $\hat{V}$ and $\tilde{V}$ have the following properties:

- $\tilde{V} < 1$
- $\hat{V} > \tilde{V}$
- $\hat{V} > 1 \text{if} \rho + \eta + \upsilon(\kappa_D) - (1-\gamma)\left(P_D - \upsilon(\kappa_D)\right) - \eta\beta(\alpha, b, \kappa, \mu^\delta, \delta, \zeta_D)^{1-\gamma} > 0$

- $\underset{V \to 0}{lim} R(V) \to -\infty$
- $\underset{V \to +\infty}{lim} R(V) \to 0$
- $R(1) = (\rho + \eta) - (1-\gamma)P_D - \eta\beta(\alpha, b, \kappa, \mu^\delta, \delta, \zeta_D)^{1-\gamma}$.

The properties of $R(V)$ and $L(V)$ above described, together with conditions imposed on the model parameters are sufficient conditions for the solution to exist and be unique. Furthermore, assuming $\zeta_D = \zeta = \bar{\zeta}$, when $\gamma > 1$ we have

$$L(1) > R(1)$$
$$(\rho + \eta) + (\gamma - 1)P - \eta\beta(\alpha, b, \kappa, \mu^\delta, \delta, \bar{\zeta})^{1-\gamma} > (\rho + \eta) + (\gamma - 1)P_D - \eta\beta(\alpha, b, \kappa, \mu^\delta, \delta, \bar{\zeta})^{1-\gamma}$$
$$(\gamma - 1)P > (\gamma - 1)P_D,$$

and since $L(V)$ and $R(V)$ are strictly increasing and strictly decreasing in V, the solution is such that $V^* < 1$.When $\gamma < 1$, $R(1) > L(1)$, $L(V)$ is strictly increasing in V. Since $\hat{V} < 1$, V is in the region of the domain where $R(V)$ it is strictly decreasing, and the solution is such that $V^* > 1$. The properties of the equation that pins down the equilibrium are summarized in the figure below:

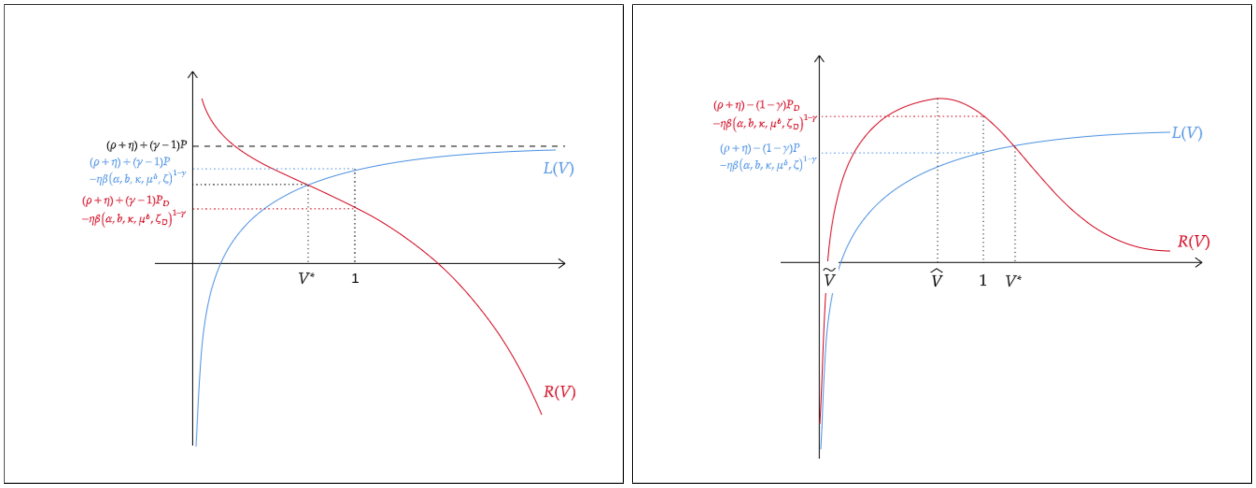
B.1.3 Proposition 2.

The proof follows from equation (B.11), which pins down the equilibrium value of V^*.

$$L(V) = R(V)$$

Figure B.1: Characterization of the Equilibrium for Different Values of Intertemporal Elasticity of Substitution.

($\textit{Left Panel}$ $\gamma > 1$ - $\textit{Right Panel}$ $\gamma \in (0, 1)$)



$$(\rho+\eta) - (1-\gamma)P - \eta V^{-1}\beta(\alpha, b, \kappa, \mu^{\delta}, \delta, \zeta)^{1-\gamma} =$$
$$\left[\rho + \eta + \upsilon(\kappa_D) - (1-\gamma)P_D - \eta\beta(\alpha, b, \kappa, \mu^{\delta}, \delta, \zeta_D)^{1-\gamma} - \upsilon(\kappa_D)V\right] V^{-1/\gamma}$$

There are two possible cases:

High intertemporal elasticity of substitution ($\gamma \in (0, 1)$).

In this case $L(V)$ is strictly increasing in V,and in a neighbourhood V^*, $R(V)$ is strictly decreasing in V.

- $\frac{\partial V^*}{\partial \zeta} > 0$

  An increase in ζ rises $\beta(\alpha, b, \kappa, \mu^{\delta}, \delta, \zeta)^{1-\gamma}$, which reduces $L(V)$. Restoring the equilibrium requires an increase in V^*.

- $\frac{\partial V^*}{\partial \zeta_D} < 0$

  An increase in ζ_D rises $\beta(\alpha, b, \kappa, \mu^{\delta}, \delta, \zeta_D)^{1-\gamma}$, which reduces $R(V)$. Restoring the equilibrium requires a reduction in V^* to reduce $L(V)$.

- $\frac{\partial V^*}{\partial \upsilon(\kappa_D)} < 0$

  $\frac{\partial R(V)}{\partial \upsilon(\kappa_D)} = V^{-1/\gamma}(1-V) < 0$. Restoring the equilibrium requires a reduction in V^* to reduce $L(V)$.

- $\frac{\partial V^*}{\partial \eta}|_{\zeta=\zeta_D} < 0$

  $\frac{\partial R(V)}{\partial \eta} = \left(1-\beta(\alpha,b,\kappa,\mu^\delta,\delta,\zeta)^{1-\gamma}V^{-1}\right) > \frac{\partial L(V)}{\partial \eta} = \left(V^{-1/\gamma}\left[1-\beta(\alpha,b,\kappa,\mu^\delta,\delta,\zeta_D)^{1-\gamma}\right]\right)$.

  Restoring the equilibrium requires a reduction in V^*.

Low intertemporal elasticity of substitution ($\gamma > 1$).

- $\frac{\partial V^*}{\partial \zeta} < 0$

  An increase in ζ reduces $\beta(\alpha,b,\kappa,\mu^\delta,\delta,\zeta)^{1-\gamma}$, which increases $L(V)$. Restoring the equilibrium requires a reduction in V^*.

- $\frac{\partial V^*}{\partial \zeta_D} > 0$

  An increase in ζ_D leads to a decline in $\beta(\alpha,b,\kappa,\mu^\delta,\delta,\zeta_D)^{1-\gamma}$, which increases $R(V)$. Restoring the equilibrium requires a rise in V^* to increase $L(V)$.

- $\frac{\partial V^*}{\partial \upsilon(\kappa_D)} > 0$

  Since $V < 1^*$,$\frac{\partial R(V)}{\partial \upsilon(\kappa_D)} = V^{-1/\gamma}(1-V) > 0$. Restoring the equilibrium requires a rise in V^* to increase $L(V)$.

- $\frac{\partial V^*}{\partial \eta}|_{\zeta=\zeta_D} < 0$

  $V < 1^* \implies$

  $\frac{\partial R(V)}{\partial \eta} = \left(1-\beta(\alpha,b,\kappa,\mu^\delta,\delta,\zeta)^{1-\gamma}V^{-1}\right) < \frac{\partial L(V)}{\partial \eta} = \left(V^{-1/\gamma}\left[1-\beta(\alpha,b,\kappa,\mu^\delta,\delta,\zeta_D)^{1-\gamma}\right]\right)$.

  Restoring the equilibrium requires an increase in V^*.

$\square$

B.1.4 Proposition 4.

It is convenient to prove Proposition 4 first before continuing with Proposition 3. From Proposition 1, the growth rates in the *normal* and *disaster* phases are

$$\mu^N = H + \alpha_N (A - H) - \kappa - b_N - V_N^{-\left(\frac{1}{\gamma}\right)},$$

$$\mu^D = H_D + \alpha_D (A_D - H_D) - \kappa_D - b_D - V_D^{-\left(\frac{1}{\gamma}\right)}.$$

Assuming $\mu^\delta(\kappa) = \mu_k^\delta(\kappa)(1+\zeta_k) = \mu_h^\delta(\kappa)(1+\zeta_h), \Delta_k(\kappa) = 0$, and $\alpha_N = \frac{A-H}{\gamma\sigma^2}$,

$$\alpha_D = \frac{A_D - H_D}{\gamma\sigma^2}, \beta(\alpha, b, \kappa, \mu^\delta, \delta, \zeta) = \mu^\delta(\kappa) + \frac{b}{\eta} + \zeta.$$

P and P_D are given by

$$P = H + \frac{1}{2}\left(\frac{A - H}{\gamma\sigma^2}\right) - \kappa,$$

and

$$P_D = H_D + \frac{1}{2}\left(\frac{A_D - H_D}{\gamma\sigma^2}\right) - \kappa_D.$$

Moreover, V_N and V_D are pinned down by

$$\gamma V_N^{-\left(\frac{1}{\gamma}\right)} = (\rho + \eta) - (1-\gamma)P + (1-\gamma)b_N^\star - \eta V^{-1}\left(\mu^\delta(\kappa) + \frac{b}{\eta} + \zeta\right)^{1-\gamma},$$

and

$$\gamma V_D^{-\left(\frac{1}{\gamma}\right)} = (\rho + \eta + \nu(\kappa_D)) - (1-\gamma)P_D + (1-\gamma)b_D^\star - \eta\left(\mu^\delta(\kappa) + \frac{b_D}{\eta} + \zeta_D\right)^{1-\gamma} - \nu(\kappa_D)\, V.$$

And the growth rates are

$$\mu^N = H + \frac{(A-H)^2}{\gamma\sigma^2} - \kappa - b_N - \frac{(\rho+\eta)}{\gamma} + \frac{(1-\gamma)}{\gamma}P - \frac{(1-\gamma)}{\gamma}b_N^\star + \frac{\eta}{\gamma}V^{-1}\left(\mu^\delta(\kappa) + \frac{b}{\eta} + \zeta\right)^{1-\gamma}$$

$$\mu^D = H_D + \frac{(A_D-H_D)^2}{\gamma\sigma^2} - \kappa_D - b_D - \frac{(\rho+\eta+\nu(\kappa_D))}{\gamma} + \frac{(1-\gamma)}{\gamma}P_D - \frac{(1-\gamma)}{\gamma}b_D^\star$$
$$+ \frac{\eta}{\gamma}\left(\mu^\delta(\kappa) + \frac{b_D}{\eta}\zeta_D\right)^{1-\gamma} + \frac{\nu(\kappa_D)}{\gamma}V,$$

which simplify to

$$\mu^N = \frac{1}{\gamma}\left[H + \frac{(\gamma+1)}{2}\left(\frac{(A-H)^2}{\gamma\sigma^2}\right) - (\rho+\eta) - \kappa - b_N + \eta V^{-1}\left(\mu^\delta(\kappa) + \frac{b}{\eta} + \zeta\right)^{1-\gamma}\right],$$

$$\mu^D = \frac{1}{\gamma}\left[H_D + \frac{(\gamma+1)}{2}\left(\frac{(A_D-H_D)^2}{\gamma\sigma^2}\right) - (\rho+\eta+\nu(\kappa_D)) - \kappa_D - b_D\right.$$
$$\left. + \eta\left(\mu^\delta(\kappa) + \frac{b}{\eta} + \zeta_D\right)^{1-\gamma} + \nu(\kappa_D)V\right].$$

Under no insurance markets for disasters $b_N = 0$ and $b_D = 0$. Focus on the case where $V > 1$.

- $\frac{\partial \mu^N}{\partial \zeta} = \eta(1-\gamma)V^{-1}\left(\mu^\delta(\kappa)+\zeta\right)^{-\gamma} - \left(\frac{\partial V}{\partial \zeta}\right)\eta V^{-2}\left(\mu^\delta(\kappa)+\zeta\right)^{1-\gamma}$
  $= \eta V^{-1}\left(\mu^\delta(\kappa)+\zeta\right)^{-\gamma}\left((1-\gamma) - \left(\frac{\partial V}{\partial \zeta}\right)\eta V^{-1}\left(\mu^\delta(\kappa)+\zeta\right)\right) > 0 \text{ or } < 0.$
- $\frac{\partial \mu^N}{\partial \zeta_D} = \eta(-1)V^{-2}\left(\frac{\partial V}{\partial \zeta_D}\right) > 0.$
- $\frac{\partial \mu^D}{\partial \zeta} = \left(\frac{1}{\gamma}\right)\nu(\kappa_D)\left(\frac{\partial V}{\partial \zeta}\right) < 0.$
- $\frac{\partial \mu^D}{\partial \zeta_D} = \left(\frac{1}{\gamma}\right)\left(\eta(1-\gamma)\left(\mu^\delta(\kappa)+\zeta_D\right)^{-\gamma} + \nu(\kappa_D)\left(\frac{\partial V}{\partial \zeta_D}\right)\right) > 0 \text{ or } < 0.$

□

B.1.5 Proposition 3.

Under full insurance, in the normal phase $V\left(\mu^{\delta}(\kappa)+b_N+\zeta\right)^{\gamma}=1$ and $\left(\mu^{\delta}(\kappa)+b_D+\zeta_D\right)=1$. As a consequence, the equation that pins down the relative valuation is independent of ζ_D, and the growth rates are given by

$$\mu^N=\frac{1}{\gamma}\left[H+\frac{(\gamma+1)}{2}\left(\frac{(A-H)^2}{\gamma\sigma^2}\right)-(\rho+\eta)-\kappa-b_N+\eta\left(\mu^{\delta}(\kappa)+\frac{b_N}{\eta}+\zeta\right)\right],$$

and

$$\mu^D=\frac{1}{\gamma}\left[H_D+\frac{(\gamma+1)}{2}\left(\frac{(A_D-H_D)^2}{\gamma\sigma^2}\right)-(\rho+\nu(\kappa_D))-\kappa_D-b_D+\nu(\kappa_D)V\right].$$

- $\frac{\partial\mu^N}{\partial\zeta}=\left(\frac{1}{\gamma}\right)\eta>0$
- $\frac{\partial\mu^N}{\partial\zeta_D}=0$
- $\frac{\partial\mu^N}{\partial\bar{\zeta}}=\left(\frac{1}{\gamma}\right)\eta>0$
- $\frac{\partial\mu^D}{\partial\zeta}=\left(\frac{1}{\gamma}\right)\nu(\kappa_D)\left(\frac{\partial V}{\partial\zeta}\right)<0$
- $\frac{\partial\mu^D}{\partial\zeta_D}=0$
- $\frac{\partial\mu^D}{\partial\bar{\zeta}}=\left(\frac{1}{\gamma}\right)\nu(\kappa_D)\left(\frac{\partial V}{\partial\zeta}\right)<0$

$\square$

B.1.6 Proposition 5.

Assume the country has full access to disaster insurance. In this case, the equation that pins that the relative valuation in equilibrium is

$$(\rho+\eta)-(1-\gamma)P-\eta\left(\mu^{\delta}(\kappa)+\frac{b_N}{\eta}+\zeta\right)=\left[\rho+\upsilon(\kappa_D)-(1-\gamma)P_D-\upsilon(\kappa_D)V\right]V^{-1/\gamma}.$$

- Changes in disaster frequency (η)

  - $\frac{\partial \mu^N}{\partial \eta} = \left(\frac{1}{\gamma}\right)\left(-1 + \left(\mu^\delta(\kappa) + \frac{b_N}{\eta} + \zeta\right)\right) < 0$

  - $\frac{\partial \mu^D}{\partial \eta} = \left(\frac{1}{\gamma}\right)\left(\nu\left(\kappa_D\right)\left(\frac{\partial V}{\partial \eta}\right)\right) \begin{cases} < 0 \text{ if } \gamma > 1 \\ > 0 \text{ if } \gamma \in (0,1) \end{cases}$

- Changes recovery speed $\nu\left(\kappa_D\right)$

  - $\frac{\partial \mu^N}{\partial \nu(\kappa_D)} = 0$

  - $\frac{\partial \mu^D}{\partial \nu(\kappa_D)} = \left(\frac{1}{\gamma}\right)(-1+V) \begin{cases} < 0 \text{ if } \gamma > 1 \\ > 0 \text{ if } \gamma \in (0,1) \end{cases}$

- Changes in prevention technologies $\mu^\delta\left(\kappa\right)$

  - $\frac{\partial \mu^N}{\partial \mu(\kappa)} = \left(\frac{1}{\gamma}\right)\eta > 0$

  - $\frac{\partial \mu^D}{\partial \mu(\kappa)} = \left(\frac{1}{\gamma}\right)\left(\nu\left(\kappa_D\right)\left(\frac{\partial V}{\partial \mu(\kappa)}\right)\right) > 0$

- Changes in productivity loss associated to disasters ϕ. Defining $\frac{P_D}{P} = 1 - \phi$

  - $\frac{\partial \mu^N}{\partial \phi} = \left(\frac{1}{\gamma}\right)(P_D(-1)) < 0$

  - $\frac{\partial \mu^D}{\partial \phi} = \left(\frac{1}{\gamma}\right)\left(-P + \nu\left(\kappa_D\right)\left(\frac{\partial V}{\partial \phi}\right)\right) < 0$

B.1.7 Proposition 6.

Assume $\gamma > 1$. First, to pin down how changes in κ affect V^*, we compute

$$\frac{\partial L(V)}{\partial \kappa} = -(1-\gamma)\frac{\partial P}{\partial \kappa} - \eta(1-\gamma)\beta(\alpha, b, \kappa, \mu^\delta, \delta, \zeta_D)^{-\gamma}\mu'\left(\kappa\right) = (1-\gamma)\left(1 - \eta\mu'\left(\kappa\right)\beta(\alpha, b, \kappa, \mu^\delta, \delta, \zeta_D)^{-\gamma}\right) < 0.$$

Thus, restoring the equilibrium requires an increase in V^* to reduce $R(V)$, and $\frac{\partial V}{\partial \kappa} > 0$.

Additionally for all γ:

$$\frac{\partial V}{\partial \kappa_D} = 0 \iff \frac{\partial R(V)}{\partial \kappa_D} = 0 \iff \nu' (\kappa_D) (1 - V) - (1 - \gamma) \left(\frac{\partial P_D}{\partial \kappa_D} \right) = 0$$

$$(1 - V) = \frac{(1 - \gamma)}{\nu' (\kappa_D)} \left(\frac{\partial P_D}{\partial \kappa_D} \right)$$