ESSAYS ON GROWTH, TRADE AND INEQUALITY

by

Gino A. Gancia

INSTITUTE FOR INTERNATIONAL ECONOMIC STUDIES
Stockholm University

Monograph Series
No. 45

2003

THE INSTITUTE FOR INTERNATIONAL ECONOMIC STUDIES
University of Stockholm

is an institute for independent scholarly research in the field of international economic relations.

DIRECTOR: Torsten Persson, Professor of Economics
DEPUTY DIRECTOR: Henrik Horn, Professor of International Economics

RESEARCH STAFF AND VISITING FELLOWS 2001/2002:

Professors
Lars Calmfors
Harry Flam
Henrik Horn
Assar Lindbeck
Mats Persson
Torsten Persson
Peter Svedberg
Lars E.O. Svensson
Fabrizio Zilibotti

Visiting Professor
Per Krusell

Research Fellows
John Hassler
Eva Nagypal
Dirk Niepelt
Kjetil Storesletten
David Strömberg
Jakob Svensson

Graduate Students
Thomas Eisensee
Giovanni Favara
Gino A. Gancia
Charlotta Groth
Bård Harstad
Mathias Herzing
Åsa Johansson
Emanuel Kohlscheen
Martin Ljunge
Conny Olovsson
Natalie Pienaar
Ulrika Stavlöt
Gisela Waisman

Visiting Researcher
José V. Rodríguez Mora

Visiting Fellows
Lilia Cavallari
  University of Rome I
Wouter Den Haan
  Univeristy of California San Diego
Matthias Döpke
  University of California Los Angeles
Wilfred Ethier
  University of Pennsylvania
Giulio Fella
  Queen Mary, University of London
Raquel Fernandez
  New York University
Po-Han Fong (Visiting Graduate Student)
  University of Rochester
Martín Gonzalez-Eiras
  Universidad de San Andrés
Fatih Guvenen
  University of Rochester
Paul Klein
  University of Western Ontario
Nils-Petter Lagerlöf
  University of Concordia
Antonio Merlo
  University of Pennsylvania
Peter Norman
  University of Wisconsin
Jan Van Ours
  University of Tilburg
Anssi Rantala (Visiting Graduate Student)
  University of Helsinki
Anna Sjögren
  IUI
Gianluca Violante
  University College London
Joseph Zeira
  Harvard University, Hebrew University
Yves Zenou
  University of Southampton

MAIL ADDRESS: **S-106 91 STOCKHOLM, SWEDEN**
STREET ADDRESS: **Universitetsvägen 10 A, 8th floor**
TELEPHONE: **+ 46 8 16 20 00,** **TELEFAX:** **+ 46 8 16 14 43**

ESSAYS ON GROWTH, TRADE AND INEQUALITY

by

Gino A. Gancia

INSTITUTE FOR INTERNATIONAL ECONOMIC STUDIES
Stockholm University

May 2003

Institute for International Economic Studies
Stockholm University

ISBN 91-7265-974-3

Printed by Akademitryck AB
Edsbruk, Sweden 2003

Doctoral Dissertation
Department of Economics
Stockholm University

ABSTRACT

This thesis consists of four papers in international macroeconomics.

Globalization, Divergence and Stagnation argues that, in a world where poor countries provide weak protection for intellectual property rights, market integration systematically shifts technical change in favor of rich nations. For this reason, free trade can increase the international income gap. At the same time, integration with countries where intellectual property rights are weakly protected can have a large adverse effect on the world growth rate. These results provide a strong rationale for global regulations, critical in a system of interdependent economies for sustaining innovation and reducing income inequality. Supportive empirical evidence is presented.

The Skill Bias of Word Trade shows that, under plausible assumptions about preferences and technology, the entire volume of world trade matters for wage inequality and therefore trade integration, even among identical countries, is likely to increase the skill premium. Further, it argues that empirical evidence of a falling relative price of skill-intensive goods can be reconciled with the fast growth of world trade and that the intersectoral mobility of capital exacerbates the effect of trade on inequality. New empirical evidence in support of the results and a quantitative assessment of the skill bias of world trade are provided.

Scale and Inequality extends the previous paper by showing with a simple two-sector general equilibrium model featuring imperfect competition, plant-level fixed costs and endogenous mark-ups that, under very general conditions, scale is skill-biased. This result provides an important link among the major explanations for the worldwide rise in wage inequality.

Geography, Migrations and Equilibrium Unemployment studies the effects of trade integration on the regional co-evolution of income, migrations and unemployment in a dynamic core-periphery model with limited labor mobility and frictions in the job matching process. The model can help explain a recently documented

empirical puzzle, i.e., the divergence of unemployment rates, together with low migrations and modest income convergence experienced by European regions over the last twenty years. By studying explicitly the transitional dynamics of the model, it also highlights a contrast between short run and long run effects of trade and policy shocks on a geographically differentiated economy.

To My Mother, Iolanda

ACKNOWLEDGMENTS

This work would not have been possible without the help of many people. First of all, I wish to thank my advisors, Torsten Persson and Fabrizio Zilibotti, whose continuous support and expert guidance were invaluable.

I owe a special gratitude to institutions in Stockholm. The Department of Economics at Stockholm University offered me the opportunity to start a very rewarding and enjoyable doctoral program and the IIES provided me the ideal research environment to pursue my ideas. I am also grateful to Stockholm University for financing my studies and to the Jan Wallander and Tom Hedelius Foundation for sponsoring my visit to MIT during the academic years 2001-2003.

Hospitality at the Economics Department of MIT greatly exceeded my expectations. I benefited enormously from the interaction with faculty and students there. In particular, I am indebted to Daron Acemoglu and Jaume Ventura for being like advisors to me and inspiring my work.

Many friends in Italy and Sweden contributed to make this thesis possible. At various stages, their encouragement has been essential and I am sincerely grateful to all of them. In particular, I would like to mention Paolo Epifani, a great coauthor and a friend.

Finally, my warmest thanks go to my family and Alessandra, for always supporting my decisions in these long years. I only regret the time this work took away from us.

TABLE OF CONTENTS

Chapter 1

Introduction

The unifying theme of this thesis is the notion that trade affects simultaneously economic prosperity and various forms of inequality: differences in living standards between nations, skill premia within countries, and regional disparities in unemployment. It is part of a broader research agenda aimed at studying how the increasingly important phenomenon of economic integration affects macroeconomic variables. This project moves from the observation that, in a system of interdependent markets, income distribution, localization of production, innovation and growth depend in a complex way on national policies and international regulations. Given the controversies on the state of the world economy, improving our understanding of these issues seems an important challenge.

In this spirit, **Chapter 2** starts by observing two stylized facts that have become the subject of a heated debate: the dramatic increase in the volume of trade between more and less developed countries, and the widening of living standard across nations in the past forty years. By studying a specific market failure common in many developing countries, this paper argues that globalization may indeed amplify income disparities. Building on the idea of directed technical change (Acemoglu, 1998, and Acemoglu and Zilibotti, 2002), it shows that North-South trade can generate divergence, through the endogenous response of technical change, if developing countries do not provide adequate protection of intellectual property rights. Since innovators cannot fully appropriate the fruits of their work in developing countries, specialization in production due to trade opening translates into a shift of R&D effort towards the activities performed in rich economies only. Therefore, trade induces "innovation diversion", making the sectors in which poor countries enjoy a comparative advantage relatively less productive. Second, and perhaps more surprisingly, the paper shows that the uneven distribution of technical progress potentially brought about by trade can also undermine the incentives to innovate, so that divergence can open the door to stagnation. Rather than raising warnings against globalization, the analysis suggests that trade liberalization in developing countries should be accompanied by reforms aimed at a tightening of intellectual property rights. With the inclusion of the TRIPS agreement in the WTO, international negotiations have recently taken important steps in that direction. A major contribution of this paper is thus to provided new theoretical foundations for these efforts. However, imposing common standards can be costly for some less developed countries and will not be sufficient to attract appropriate technologies if the economic weight of the South is low. Given these distortions, promoting research aimed at the needs of less developed countries appears to be a key element for reducing cross-country income differences and fostering world growth.

Chapter 3 tackles the effects of trade on within-country income inequality. It is a well-documented fact that wage inequality widened over the last two decades, particularly in OECD countries. Among the major explanations put forward is the growing competition with imports from low-wage producers. However, this simple trade story is inconsistent with three facts: volumes of North-South trade are still too small to have a large impact on wages, skill premia rose also in many developing countries, whereas the theory predicts the opposite, and the relative price of skill-intensive goods did not increase during the period of rising inequality. Despite these failures, many economists still believe that globalization is an important factor in the dynamics of inequality. Moving from these observations, chapter 3 proposes a new mechanism through which international trade increases wage inequality compatible with the empirical evidence. It does so by revisiting the new trade theory's account of the distributional effects of intra-industry trade (Helpman and Krugman, 1985). The main result of the chapter is to show that, if goods are gross substitutes and scale economies are higher in the skill-intensive sectors, assumptions that find empirical support, any increase in the volume of trade, even between identical countries, tends to be skill-biased. This suggests that the entire volume of world trade matters for inequality, not only the small volumes of North-South trade. Further, if the skill-biased scale effect is strong enough, this mechanism can spur inequality even in the skill-poor developing economies. Finally, it can explain the observed decline in the relative price of skill-intensive goods during the period of rising skill premia. The analysis ends with an empirical exercise. Consistently with the theory, estimations from a panel of countries in the years 1970, 1980 and 1990 show that measures of market size are positively correlated with skill premia: a doubling of the degree of trade openness is found to be associated with a 30% increase in the skill premium and a doubling of country size with a 7% increase in the skill premium.

Chapter 4 pushes the previous analysis further in important dimensions. First, it adds micro-foundations for the sectoral asymmetry in increasing returns that makes scale skill biased. Second, it argues that the resulting mechanism is so general to provide a link among the major explanations for the increase in wage inequality: skill-biased technical change, capital-skill complementarity and international trade. Since all these phenomena imply an expansion of market size, their scale effect naturally generates an upward pressure on the skill premium, even in the absence of technology biases, complementarities among inputs or Stolper-Samuelson effects. To accomplish this, the model in Chapter 3 is modified by making the degree of competition endogenous in each sector. When firm are allowed to compete a la Cournot, a scale expansion involves a pro-competitive effect that forces firms to lower mark-ups and expand output. This effect, generating increasing returns to scale, is stronger the lower the starting level of competition in a sector. Therefore, as long as skill-intensive sector is less competitive than the other, any increase in market size (due, for instance, to factor augmenting technical progress, factor accumulation or trade integration) brings about an increase in the skill premium. Interestingly, even in the absence of any asymmetries in technology, this will be the case if the skilled workers represent a minority of the labor force (true, almost by definition). More in general, the skill bias of scale turns out to be larger the lower the share of skilled workers in the total workforce, the greater the relative importance of plant-level fixed costs and the lower the relative degree of substitutability among skill-intensive goods.

Chapter 5 is focused on market integration and inter-regional inequality. In the last decades, Western Europe has undergone a process of deep economic integration. Recent developments in the field of the new economic geography (Fujita et al., 1999) have shown that such a process may trigger the spatial agglomeration of economic activity. However, this literature neglects imperfections in the labor market and hence it cannot explain the geography of unemployment. Yet, the evidence concerning European regions shows a strong tendency toward polarization and divergence of regional unemployment rates, together with a slight tendency toward convergence in per capita income. As a consequence, the uneven spatial distribution of unemployment is nowadays the main cause of policy concern in Europe. To study the geography of unemployment, Chapter 5 introduces imperfections in the job matching process (Pissarides, 1990) in a core-periphery model (Krugman, 1991) with limited labor mobility. The resulting model is then used to study the effect of trade integration on income inequality, unemployment differentials and migration flows. The main findings are the following. Starting from a symmetric equilibrium with high trade costs, trade integration triggers a wave of migrations which lead to the emergence of a core-periphery equilibrium, with strong regional disparities both in terms of per capita income and unemployment rates. Thereafter, a further reduction of trade barriers generates income convergence, which in turn reduces the incentive to migrate. At the same time, the unemployment differential still grows larger until a higher level of integration is reached. Assuming that in the early Eighties the European regions were in a core-periphery equilibrium, the model implies that regional integration generates a pattern broadly consistent with the evidence: convergence in real per capita income, divergence in regional unemployment rates, and declining migration rates. A contribution of the paper is also to highlight a sharp contrast between the short run impact of labor mobility and its effects on the long run equilibrium. In particular, during the transitional dynamics migration tends to promote regional unemployment convergence, whereas in the long run it exacerbates the unemployment differential. Similarly, changes in a common policy implemented by a central government can affect in a very different way the two regions, both in the short run dynamics and in the long run.

References

Acemoglu, Daron and Fabrizio Zilibotti (2001). "Productivity Differences," *Quarterly Journal of Economics* 116, 563-606.

Acemoglu, Daron (1998). "Why Do New Technologies Complement Skills? Directed Technical Change and Wage Inequality," *Quarterly Journal of Economics* 113, 1055-1090.

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

Helpman, Elhanan and Paul Krugman (1985). *Market Structure and Foreign Trade*, MIT Press, Cambridge.

Krugman, Paul (1991). "Increasing Returns and Economic Geography," *Journal of Political Economy* 99, 483-499.

Pissarides, Christopher (1990). *Equilibrium Unemployment Theory*, Basil Blackwell, Oxford.

Chapter 2

Globalization, Divergence and Stagnation*

1 Introduction

The past decades have witnessed a dramatic increase in the level of market integration across the globe. During the period 1960-1998, the average share of import plus export in total GDP rose from 0.54 to 0.76 and the total volume of world merchandize trade grew steadily at 10.7% per year.[1] A distinctive feature of this wave of globalization is the increasingly important role played by less developed countries (LDCs). Although trade between the US and non-OECD countries is still relatively small, it almost tripled during the period 1980-95 (Wood, 1998) and the same years have seen unprecedented episodes of market liberalization in LDCs (Sachs and Warner, 1995). In this scenario of increasing integration between more and less advanced economies, the cross-country income distribution is also changing. Many commentators claim that we live in an era of growing inequality. Quah (1993)

* I am very grateful to Daron Acemoglu, Torsten Persson, Jaume Ventura and Fabrizio Zilibotti for many insightful discussions. I also thank Pol Antras, Francesco Caselli, Paolo Epifani, Diego Puga, Paul Segerstrom, Bob Staiger, Dan Trefler and seminar participants at MIT, IIES, Stockholm University, Stockholm School of Economics, University of British Columbia, Toronto, Wisconsin, Rochester, CREI, Pompeu Fabra, UCL, Bocconi University and the European Winter Meeting of the Econometric Society (Budapest, 2002) for helpful comments. The usual caveat applies. Financial support from the Wallander and Hedelius Foundation is gratefully acknowledged. I thank the MIT Economics Department for hospitality.

[1] The trade share in GDP is from the Penn World Table, Mark 6.0; averages refer to a constant sample of 115 countries. World merchandize trade is from WTO data.

documents that countries are diverging from the world mean.[2] Similarly, Pritchett (1997) argues that "divergence in relative productivity levels is the dominant feature of modern economic history".[3] Despite evidence of convergence among rich nations and falling poverty in world population,[4] a crude measure of cross-country inequality, the variance of log real per capita GDP, displays a disturbing upward trend, rising steadily from 0.7 in 1960 to more than 1.3 in 1998.[5] Observations like these stress the centrality of understanding the effects of trade on the world income distribution and raise the concern of a possible causal link from globalization to divergence. This concerns have recently been the subject of heated debates. Although it is well known that trade affects the world income distribution, only few models focus on how and why gains from trade may be systematically biased in favor of rich nations.[6]

By studying a specific market failure common in many developing countries, this paper argues that globalization may indeed amplify income disparities. First, it shows that North-South trade can generate divergence, through the endogenous response of technical change, if developing countries do not provide adequate protection of intellectual property rights (IPRs). Since innovators cannot fully appropriate the fruits of their work in developing countries, specialization in production due to trade opening translates into a shift of R&D effort towards the activities performed in rich economies only. Therefore, trade induces "innovation diversion", making the sectors in which poor countries enjoy a comparative advantage relatively less productive. Second, the paper shows that the uneven distribution of technical progress potentially brought about by trade can also undermine the incentives to innovate,

---

[2]Interestingly, Beaudry, Collard and David (2002) show that this phenomenon seems to be more pronounced among open countries.

[3]Pritchett (1997), using data from Maddison (1995), shows that, over the past century, advanced economies consistently grew faster than the less developed ones. Perhaps surprisingly, the average growth differential reaches a peak in the last two decades, characterized not olny by the globalization boom, but also by low productivity growth in advanced countries.

[4]See Sala-i-Martin (2002) on falling poverty in world population, a phenomenon mainly due to the good performance of two very populous countries, India and China. For the purpose of the paper, that is to relate different policies to economic performance, the country seems the relevant unit of analysis. See Acemoglu and Ventura (2003) on the relative stability of the world income distribution.

[5]Data form the Penn World Table 6.0 on a sample of 115 countries.

[6]The most common argument is based on the need to protect infant industry in LDCs. See Young (1991) and Acemoglu, Aghion and Zilibotti (2002) for recent applications.

so that divergence can open the door to stagnation.

To make this argument, the paper builds a Ricardian model with endogenous, sector specific, technical change. Two sets of countries, the North and the South, are distinguished by exogenous sectoral productivity differences. Except for this Ricardian element, defining the pattern of comparative advantage, countries have access to the same pool of technologies, whose productivity can be increased by innovation. Innovation is financed by the rents it generates, but in the South some rents are dissipated due to imitation. The model is solved under autarky and free trade and the two equilibria are compared. In both cases, the equilibrium has a number of desirable properties: the world income distribution is stable, growth rates are equalized across sectors, countries with higher exogenous productivity levels are relatively richer. But the world income distribution depends crucially on the trade regime. With no commodity trade, each country produces the whole range of goods and therefore each innovator, serving the world economy, obtains both the high rents from the North and the smaller rents form the South. Under free trade, instead, each country specializes in the sectors where it has a comparative advantage and innovators obtain the rents from one location only. Since the rents from the South are smaller, the Southern sectors attract less innovation which, over time, reduces their productivity. This is the first result of the paper: in a world where poor countries provide weak protection for IPRs, market integration shifts technical change in favor of the rich ones.

Is then North-South trade always beneficial for advanced countries? The somehow surprising answer, leading to the second result of the paper, is not necessarily: under free trade, weak IPRs have a strong potential to disrupt incentives for innovation and growth, thereby hurting all countries. As the North becomes relatively richer, more sectors move to the South, where production costs are lower, and R&D becomes less attractive for a wider range of goods. Divergence can thus be followed by stagnation. In the limit case of no IPRs protection at all in the South, this process generates decreasing returns to innovation and growth eventually stops. Therefore, the model shows that in a world of interdependent economies, the regulatory policies of each country are crucial to sustain the growth rate of the entire global system.

These results have important implications. First, they provide strong arguments in favor of global protection of IPRs. In an era of falling trade barriers and increasing internationalization of production, the enforcement of IPRs in all parts of the world becomes critical for attracting and sustaining innovation. Second, that the desirability of IPRs depends on the trade regime can shed light on an observed change in attitudes of more and less advanced countries towards protection of intellectual property. The importance of defining common regulations in a global economy was recognized by the inclusion of the Agreement on Trade Related Intellectual Property Rights in the statute of the WTO.[7] As the relocation of production in less developed countries can undermine growth in the entire system, rich economies have indeed a strong incentive to put pressure for a tightening of global regulations. Similarly, less advanced countries appear more willing to provide protection for IPRs in exchange for a better access to international markets. In this respect, this paper is the first to provide a rationale for linking trade liberalization to a tightening of IPRs and suggests that the TRIPS agreement, despite the criticism of the skeptics, may actually alleviate some undesirable distributional effects of globalization. Third, contrary to the view of industrial-policy advocates, suggesting that developing countries should try to target high growth sectors, the model warn that any sector can become stagnant if incentives to innovation become weak and that industrial targeting can be less effective than hoped.

The results of the paper are based on four assumptions: specialization driven by trade, sector-specific technical progress, imperfect appropriability of profits from innovation in developing countries and an elasticity of substitution between goods higher than one. All of them seem plausible and are shared by many models. That countries specialize in different sets of products, at least to some extent, appears reasonable. More specifically, the Ricardian model has proven to be useful in the literature on trade and technology and the absence of factor price equalization makes it suitable for analyzing the world income distribution. Several observations suggest that technical progress has a strong sectoral dimension. For example, R&D is mainly performed by large companies and therefore directed to their range of activities. Although innovation certainly generates spillovers, Jaffe et al. (1993) show

[7]The TRIPS agreement establishes minimum standards of protection for several categories of IPRs and a schedule for developing countries to adopt them.

that these are generally limited to products in similar technological categories.[8] Infringements of IPRs in developing countries is indeed a significant phenomenon, as proven by the many complaints of large companies based in industrial countries. In this respect, the US Chamber of Commerce estimated a profit loss for US firms of about $24 billion in 1988. Finally, gross substitutability between goods seem realistic, as it yields the sensible prediction that fast growing sectors and countries become relatively richer.

The paper is related to the vast literature on endogenous growth and trade. The model with the closest setup to the present is perhaps the one suggested by Taylor (1994), who studies growth, IPRs and trade in a Ricardian model with sector-specific innovation. However, the assumption of a unit elasticity of substitution between goods prevents him from studying distributional issues related to sectoral growth. Acemoglu and Ventura (2002) study how trade generates a stable world income distribution, but they do not analyze IPRs, innovation and imitation. Acemoglu and Zilibotti (2001) focus on factor-specific technical progress in a model where developing countries do not protect IPRs and show how this leads to the development of technologies not appropriate for the skill-endowment of the South. Despite the similar set-up, in their model trade has quite different implications, as it generates productivity convergence and leaves the world growth rate unaffected.[9] Further, in their model the world growth rate is independent on the trade regime. Closer to the spirit of the earlier endogenous growth approach, Young (1991) builds a model of learning by doing where trade can slow down the growth rate of a country that specializes in a sector with weak dynamic scale economies. The result of this paper is more general, as it shows that trade induces innovation diversion in favor of rich countries irrespective of the sector of specialization, because what matters for attracting innovation is not a characteristic of sectors, but an institutional feature of countries.

The paper is also related to the formal literature on IPRs, imitation and wel-

[8] Cross-sectoral spillovers can be included in the model without affecting the qualitative results as long as spillovers are less beneficial than a targeted innovation.

[9] Acemolgu and Zilibotti (2001) claim, without proving it, that trade, by inducing skill-biased technical change, increases the North-South income gap. It turns out that this result holds only under special circumstances. What is general, in their model, is that the endogenous response of technology makes trade less beneficial for poor countries than would othewise be.

fare, that goes back to the product cycle Ricardian model of Krugman (1979). A number of papers used his approach to study several aspects of the issue, including the effects of licensing or FDI. The earlier contributions highlighted the negative effects of strong IPRs as they would restrict the efficient allocation of resources.[10] More recently, the view that IPRs can foster growth and stimulate the diffusion of technology has gained more consensus.[11] Abstracting from product cycles, this paper offers a complementary view based on cost-saving innovations that yields new results in favor of IPRs protection. An important virtue of this approach is that it incorporates the idea that technologies can be inappropriate for developing countries and that IPRs protection can play a role in attracting better technologies. These important considerations are absent in most of the product-cycle literature.[12] Further, these models do not usually deal with the effects of IPRs under different trade regimes. Another strand of literature focuses on the welfare effects of the monopoly distortion introduced by patent laws in a trading environment.[13] In comparison, this paper shows that asymmetric regulations generate a new inefficiency, innovation diversion, that should be taken into account in designing an optimal system of international protection of intellectual property.

Finally, this analysis is complementary to Matsuyama (2000). He develops a Ricardian model where the North has a comparative advantage in high income elasticity goods. In his set up, a uniform and exogenous increase of world productivity results in a terms-of-trade deterioration for the South, because it raises the demand for the good in which the North has a comparative advantage. But Matsuyama's paper does not study the effects of the trade on technical progress, which is the main theme here.

The rest of the paper is organized as follows. Section 2 presents the basic two-country model, solves for the equilibrium under autarky and free trade and derives the two main results, that trade integration with a country where IPRs are weak can lead to divergence in income levels and slow down world growth. The analysis

[10]Among these models are Helpman (1993) and Glass and Saggi (1995) and more recently, Dinopoulos and Segerstrom (2003).

[11]Among these model, see Lai (1998), Yang and Maskus (2001) and Antras (2002).

[12]See, for example, Kremer (2002), Sachs (1999), Diwan and Rodrik (1991), and Acemoglu and Zilibotti (2002).

[13]See Chin and Grossman (1990), Deardorff (1992) and recently Grossman and Lai (2002).

ends with an extension of the results to a multi-country world and one where trade is costly. Section 3 shows some supportive empirical evidence. Section 4 concludes.

2 The Model

2.1 Autarky

Consider first the set N of rich countries (the North). The North is assumed to be a collection of perfectly integrated economies with similar characteristics, whose total population is L_N. The subscript N is suppressed where it causes no confusion. Consumers have identical isoelastic preferences:

$$U = \int_0^\infty \ln c(t) \, e^{-\rho t} dt.$$

There is a continuum $[0,1]$ of sectors, indexed by i. Output of each sector, $y(i)$, is aggregated in bundle Y used both for consumption and investment:

$$Y = \left[\int_0^1 y(i)^{\frac{\epsilon-1}{\epsilon}} di \right]^{\frac{\epsilon}{\epsilon-1}}, \tag{2.1}$$

where $\epsilon > 1$ is the elasticity of substitution between any two goods. The relative demand obtained by maximizing (2.1) is:

$$\frac{p(i)}{p(j)} = \left[\frac{y(i)}{y(j)} \right]^{-1/\epsilon}. \tag{2.2}$$

The aggregate Y is taken as the numeraire and its price index is therefore set equal to one:

$$P = \left[\int_0^1 p(i)^{1-\epsilon} di \right]^{\frac{1}{1-\epsilon}} = 1. \tag{2.3}$$

Each good $y(i)$ is homogeneous and produced by competitive firms using machines $x(i)$ and labor $l(i)$:

$$y(i) = A(i)^\beta x(i)^{1-\beta} l(i)^\beta, \tag{2.4}$$

where $A(i)$ is an index of machine productivity in sector i. Machines are sector-specific, non tradeable and depreciate fully after use. Demand for machine $x(i)$

derived from (2.4) is:

$$x(i) = [(1-\beta)\, p(i) / \chi(i)]^{1/\beta} A(i)\, l(i)\,, \tag{2.5}$$

where $\chi(i)$ is the price of machine $x(i)$. Machines in each sector are produced by a monopolist. The unit cost of producing any machine is normalized to $(1-\beta)^2$. Together with isoelastic demand (2.2), this implies that the monopolist in each sector charges a constant price, $\chi(i) = (1-\beta)$. Substituting $\chi(i)$ and (2.5) into (2.4), yields the quantity produced in sector i as a linear function of the level of technology $A(i)$ and employed labor $l(i)$:

$$y(i) = p(i)^{(1-\beta)/\beta} A(i)\, l(i)\,. \tag{2.6}$$

The linearity of $y(i)$ in $A(i)$ is crucial for endogenous growth, but it is not a sufficient condition. As it will become clear later on, an expansion of $y(i)$ can reduce its price $p(i)$ and this can effectively generate decreasing returns. Given the Cobb-Douglas specification in (2.4), the wage bill in each sector is a fraction β of sectoral output. Therefore, equation (2.6) can be used to find the relation between equilibrium prices and the wage:

$$w = \beta p(i)^{1/\beta} A(i)\,. \tag{2.7}$$

Since there is perfect mobility of labor across sectors, the wage rate has to be equalized in the economy. Dividing equation (2.7) by its counterpart in sector j delivers the equilibrium relative price of any two varieties:

$$\frac{p(i)}{p(j)} = \left[\frac{A(j)}{A(i)}\right]^{\beta}. \tag{2.8}$$

Intuitively, sectors with higher productivity have lower prices. Using (2.7), integrating over the interval $[0,1]$ and making use of (2.3) shows that the equilibrium wage rate is a CES function of sectoral productivity:

$$w = \beta \left[\int_0^1 A(i)^{\beta(\epsilon-1)}\, di\right]^{1/\beta(\epsilon-1)}. \tag{2.9}$$

Using (2.6) and (2.8) in (2.2) yields the optimal allocation of workers across sectors. Integrating over the interval $[0, 1]$ gives:

$$l(i) = L \frac{A(i)^{\beta(\epsilon-1)}}{\int_0^1 A(j)^{\beta(\epsilon-1)} dj}. \tag{2.10}$$

Note that more productive sectors attract more workers (as long as $\epsilon > 1$) because the value of marginal productivity of labor has to be equalized. Profits generated by the sale of machine i are a fraction $\beta(1-\beta)$ of the value of sectoral output:

$$\pi(i) = \beta(1-\beta) p(i)^{1/\beta} A(i) l(i). \tag{2.11}$$

The evolution of technology combines Ricardian elements with endogenous technical change. The productivity index $A(i)$ in each sector is the product of two components, an exogenously given productivity parameter, $\phi(i)$, and the level of current technology in use in sector i, $a(i)$:

$$A(i) = a(i) \phi(i).$$

While $\phi(i)$ is fixed and determined by purely exogenous factors, such as the specific environment of a country, $a(i)$ can be increased by technical progress. For simplicity, the model assumes that all the countries in the North share the same productivity schedule $\phi = (\phi(i))$. Innovation is directed and sector specific. To simplify, without loss of generality, innovation is modelled as incremental:[14] in the R&D sector, μ units of the numeraire can increase the productivity of machine i by $\partial a(i)$. Once an innovation is discovered, the innovator is granted a perpetual monopoly over its use. The patent is then sold to the producer of machine i. Free-entry in the R&D sector drives the price of any innovation down to its marginal cost μ. The monopolist decides how much innovation to buy by equating the marginal value of the quality improvement, the present discounted value of the infinite stream of profits generated by the innovation, to its cost. Along the balanced growth path, where $\partial\pi(i)/\partial a(i)$

[14]This description of innovation is equivalent to the expanding variety approach of Romer (1990). See Gancia and Zilibotti (2003) for more details on growth through expanding variety of intermediates and how to rewrite the present model in that context.

and r are constant, this condition is:

$$\frac{\partial \pi (i)}{\partial a (i)} \frac{1}{r} = \mu.$$

Using (2.11), (2.10), (2.7) and normalizing $\mu = \sigma (1 - \beta) \beta$, the previous expression reduces to:[15]

$$L\phi (i) \left[\frac{\beta w}{A (i)} \right]^{1-\beta(\epsilon-1)} = r. \tag{2.12}$$

For the remainder of the paper, define $\sigma \equiv \beta (\epsilon - 1)$ and assume $\sigma \in (0, 1)$. On the one hand, the assumption $\sigma > 0$ (equivalent to $\epsilon > 1$) rules out Bahgwati (1958) immiserizing growth: the fact that a sector (later on a country) growing faster than the others would become poorer. On the other hand, the restriction $\sigma < 1$ is required to have a stable income distribution across sectors: it implies that if a sector grows more than another, its relative profitability would fall, discouraging further innovation.[16] If violated, it would be profitable to innovate in one sector only and all the other sectors would disappear, a case that does not seem realistic. From this discussion, it is clear that along the balanced growth path R&D is performed for all the machines and all the sectors grow at the same rate. But for this to be the case, the incentive to innovate has to be equalized across sectors. Therefore, imposing condition (2.12) for all i, it is possible to characterize the equilibrium profile of relative productivity across sectors:

$$\frac{A (i)}{A (j)} = \frac{a (i) \phi (i)}{a (j) \phi (j)} = \left[\frac{\phi (i)}{\phi (j)} \right]^{\frac{1}{1-\sigma}}. \tag{2.13}$$

Equation (2.13) shows that, as long as $\sigma > 0$ (i.e., $\epsilon > 1$), sector specific innovations amplify the exogenously given productivity differences $\phi (i) / \phi (j)$. As for labor mobility, in order to equalize the returns to innovation, the exogenously more productive sectors need to have an higher than average $a(i)$.

Finally, using (2.12), (2.9) and the Euler equation for consumption growth $g =$

[15] This normalization, where σ is defined below as $\beta (\epsilon - 1)$, is meant to simplify the algebra only.

[16] Stability can be violated because the market size for innovation is proportional to $l(i)$ which is a positive function of innovation itself.

$r-\rho$, the autarky growth rate of the economy can be found as:

$$g = L\left[\int_0^1 \phi\left(i\right)^{\sigma/(1-\sigma)}\, di\right]^{(1-\sigma)/\sigma} - \rho \tag{2.14}$$

Consider now the set S of poor countries (the South). In the aggregate, the South is assumed to have a schedule of exogenously given productivity, ϕ_S, different from that of the North, ϕ_N. This Ricardian element captures the fact that geographic, cultural and economic differences (taken as exogenous) make the South relatively more advantaged in some activities compared to the North, even when technological knowledge is common. Following Dornbusch et al. (1977), sectors are conveniently ordered in such a way that the index $i \in [0,1]$ is decreasing in the comparative advantage of the North, i.e., $\phi_N\left(i\right)/\phi_S\left(i\right) > \phi_N\left(j\right)/\phi_S\left(j\right)$ if and only if $i < j$. To further simplify the analysis, assume that $\phi_N\left(i\right)$ is weakly decreasing in i and $\phi_S\left(i\right)$ is weakly increasing in i, so that the most productive sector in the North is the least productive in the South. To start with, consider the case of no protection of IPRs in the South. Still, the South is allowed to imitate at a small cost the innovations introduced in the North, so that the endogenous component of technology, $a(i)$, is identical in all the countries. This assumption reflects the quasi public good nature of technical progress, according to which only IPRs protection can exclude others from exploiting past discoveries. For simplicity, the analysis adopts a stylized description of the R&D sector in which innovators produce for the world economy and the cross-country distribution of the R&D cost is proportional to the net revenue generated from innovation in each country.[17] With no IPRs protection in the South and no trade, the Northern equilibrium is unaffected by other countries. In particular, the sectoral distribution of technical progress, $a(i)$, is determined by (2.13) according to the exogenous productivity index of the North, $\phi_N\left(i\right)$. The only difference in the South is that technical progress, embedded in $a(i)$, is taken as given from the North.[18] Using equations (2.9) and (2.13) yields the North-South wage ratio, $\omega \equiv$

[17] This assumption makes the localization of R&D irrelevant for the purpose of the analysis. Equivalently, the localization of R&D could be studied by allowing profit transfers between countries in terms of Y. In any case, given the small size of the R&D sector, about 2% of GDP in advanced countries and much less in the rest of the world, this simplification seems innocuous.

[18] In the South, each machine i will be produced by a monopolist, as in the North. In presence of a small imitation cost, no two firms have an incentive to produce the same machine because price

w_N/w_S:

$$\omega = \left[\frac{\int_0^1 \phi_N(i)^{\sigma/(1-\sigma)}\,di}{\int_0^1 \phi_N(i)^{\sigma^2/(1-\sigma)}\,\phi_S(i)^{\sigma}\,di}\right]^{1/\sigma} \tag{2.15}$$

First, note that $\partial\omega/\partial\phi_N(i) > 0$ and $\partial\omega/\partial\phi_S(i) < 0$. Intuitively, the relative wage is proportional to the exogenous productivity of the two regions, ϕ_N and ϕ_S. More important, the Appendix shows that the sectoral profile of technology is optimal for the North, in the sense that it maximizes Y_N, and is appropriate for the South only in the limit case when the two regions have the same sectoral distribution of ϕ ($\phi_S(i) = \alpha\phi_N(i), \forall i$, with α equal to a constant of proportionality).[19] This outcome mirrors, in a different set-up, the result of Acemoglu and Zilibotti (2001). Further, the Appendix shows that $\forall\sigma \in (0,1)$ ω is bounded by $\max\{\phi_N(i)/\phi_S(i)\} = \phi_N(0)/\phi_S(0)$. Lastly, since growth is due to the expansion of the $a(i)$ that are identical across countries, equation (2.14) for the North gives also the growth rate of the South.

Consider now the case of imperfect protection of IPRs in the South. To keep the analysis a simple as possible, assume that the owner of a patent can extract only a fraction θ of the profits generated by its patent in the South.[20] Therefore, θ can be interpreted as an index of the strength of IPRs protection. The profitability of an innovation is now the sum of the rents generated both in the North and in the South, and the marginal condition for buying innovations becomes:

$$\left[\frac{\partial\pi_N(i)}{\partial a(i)} + \theta\frac{\partial\pi_S(i)}{\partial a(i)}\right]\frac{1}{r} = \mu$$

competition would lead them to negative profits. The postulated independence of the monopoly distortion in the imitating South is dictated by simplicity and precludes the analysis of the trade-off between the dynamic loss and the static benefit of weak IPRs in poor countries. This trade-off, studied extensively in the literature, is particularly important for welfare analysis, which is not the main concern of the paper. On the contrary, positive rents from innovation in the South are crucial to study the case of partial protection of IPRs. This latter case seems realistic, since companies do receive royalties from developing countries.

[19] Remember that it is optimal to have high quality machines in sectors where the exogenous productivity is already high. Copying the technology from the North, the South is using high quality machines in sectors that are originally not productive. This inefficiency lowers the wage in the South.

[20] This description of IPRs is both simple and general. It can also capture practices such as licensing, where rent sharing is necessary to deter default or imitation on behalf of the licensee. See Yang and Maskus (2001) on this.

Substituting the expressions for profits and solving for $a(i)$ yields:

$$a(i) = \left[\frac{L_N \phi_N(i)^\sigma (w_N)^{1-\sigma} + \theta L_S \phi_S(i)^\sigma (w_S)^{1-\sigma}}{r}\right]^{1/(1-\sigma)} \tag{2.16}$$

Note that the endogenous component of sectoral productivity is now proportional to a weighted average of the two exogenous indexes $\phi_N(i)$ and $\phi_S(i)$, with weights that depend on country size, the strength of property rights and relative income. The general expression for the relative Northern wage becomes:

$$\omega = \left\{\frac{\int_0^1 \phi_N(i)^\sigma \left[L_N \phi_N(i)^\sigma + \theta L_S \phi_S(i)^\sigma (\omega)^{\sigma-1}\right]^{\sigma/(1-\sigma)} di}{\int_0^1 \phi_S(i)^\sigma \left[L_N \phi_N(i)^\sigma + \theta L_S \phi_S(i)^\sigma (\omega)^{\sigma-1}\right]^{\sigma/(1-\sigma)} di}\right\}^{1/\sigma} \tag{2.17}$$

Whether technology is closer to the Northern or Southern optimum, depends on which of the two markets for innovations, L_N and θL_S, is larger (see also the Appendix). As $\theta L_S/L_N \to 0$, equations (2.17) reduces to (2.15). Therefore, the case of no IPRs protection defines an upper bound for ω in autarky.

Finally, using (2.16), (2.9) and the Euler equation $g = r - \rho$, the growth rate of the world economy for the general case when $\theta \neq 0$ can be found as:

$$g = \left\{\int_0^1 \left[L_N \phi_N(i) + \theta L_S \phi_S(i)^\sigma (\phi_N(i)/\omega)^{1-\sigma}\right]^{\sigma/(1-\sigma)} di\right\}^{(1-\sigma)/\sigma} - \rho \tag{2.18}$$

Note that the world growth rate increases with θ because stronger IPRs translate into higher profits for innovation. As $\theta \to 0$, the growth rate declines to (2.14), defining a lower bound for the growth rate in autarky.

2.2 Trading Equilibrium

Trade takes place because of the Ricardian element of the model: even if technological progress is endogenous, productivity differences across countries are completely exogenous and so is comparative advantage. Recall that the ordering of sectors $i \in [0, 1]$ is decreasing in the comparative advantage of the North, so that $\phi_N(i)/\phi_S(i) > \phi_N(j)\phi_S(j)$ if and only if $i < j$. Further, for analytical tractability, the comparative advantage schedule, i.e., the ratio of exogenous productivity

$\phi_N(i)/\phi_S(i)$, is assumed to be continuous. The static equilibrium under free trade can be found imposing two conditions. The first is that each good is produced only in the country where it would have a lower price. Therefore, the North specializes in the sectors $[0, z]$ where its comparative advantage is stronger and the South produces the remaining range of goods $[z, 1]$. Given the continuity assumption on the comparative advantage schedule, the North and the South must be equally good at producing the cut-off commodity z: $p_N(z) = p_S(z)$. Using (2.7), this latter condition identifies the cut-off sector z as a function of the relative wage under free trade ω:

$$\frac{\phi_N(z)}{\phi_S(z)} = \omega. \tag{2.19}$$

Since comparative advantage of the North is decreasing in z, condition (2.19) traces a downward sloping curve, Φ, in the space (z, ω). The second equilibrium condition is trade balance, i.e., imports and exports have to be equal in value. Since total output in a country is proportional to the wage bill and the share of consumption allocated to a set $[0, z]$ of goods is $\int_0^z p(i)^{1-\epsilon} di$, trade balance can be written as:

$$w_N L_N \int_z^1 p(i)^{1-\epsilon} di = w_S L_S \int_0^z p(i)^{1-\epsilon} di$$

Note that, by homogenous tastes, the origin of demand (and R&D spending) is irrelevant. Using (2.7) the trade balance condition can be rewritten as:

$$w_N^{1+\sigma} L_N \int_z^1 A(i)^\sigma di = w_S^{1+\sigma} L_S \int_0^z A(i)^\sigma di \tag{2.20}$$

Along a balanced growth path, the profits generated by innovation in any pair of sectors must be equal. In particular, considering innovations for the Northern and the Southern markets, i and j, the following condition must hold: $\partial\pi_N(i)/\partial a(i) = \theta\partial\pi_S(j)/\partial a(j)$. Substituting (2.11) for profits, noting that under free trade the optimal allocation of labor (2.10) is $l_N(i) = L_N A_N(i)^\sigma / \int_0^z A_N(v)^\sigma dv$ and $l_S(j) = L_S A_S(j)^\sigma / \int_z^1 A_S(v)^\sigma dv$ and using (2.20), yields the equilibrium sectoral productivity profile:

$$\frac{A_N(i)}{A_S(j)} = \left[\frac{\phi_N(i)}{\theta\phi_S(j)}\right]^{1/(1-\sigma)} (\omega)^{\sigma/(\sigma-1)} \quad \forall i, j \in [0, 1] \text{ with } i \leq z \leq j \tag{2.21}$$

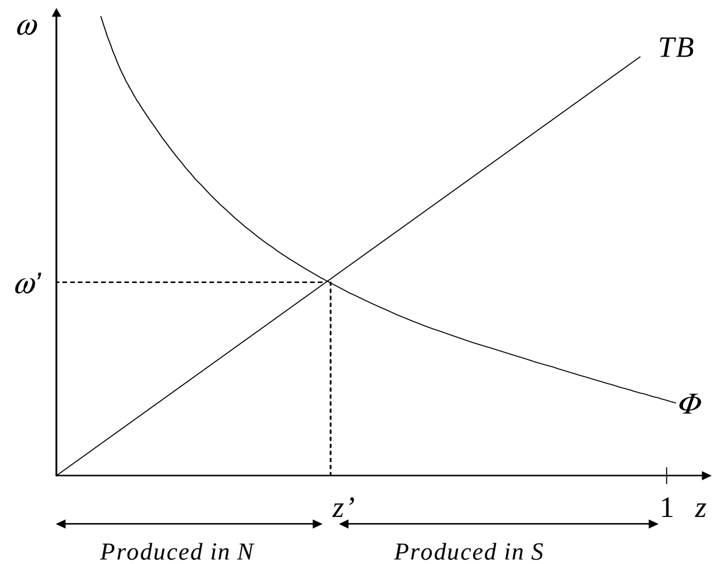

Figure 2.1: Free Trade Equilibrium

Compared to the autarky case, the relative productivity of sectors under free trade still depends on the exogenous $\phi(i)$, but also on the IPRs regime of the country where the innovation is sold. Technology is still biased towards the exogenously more productive sectors (as $\sigma \in (0,1)$, original differences $\phi_N(i)/\phi_S(j)$ are amplified) but also against the Southern sectors where some rents from innovation are lost ($\theta < 1$). Integrating i over $[0,z]$ and j over $[z,1]$ in (2.21) and using (2.20), the trade balance condition (TB), incorporating equilibrium technologies, can be rewritten as:

$$\omega = \theta^{-\sigma}\left[\frac{L_S}{L_N}\frac{\int_0^z \phi_N(i)^{\sigma/(1-\sigma)}\,di}{\int_z^1 \phi_S(i)^{\sigma/(1-\sigma)}\,di}\right]^{1-\sigma} \tag{2.22}$$

Note that ω is increasing in z and decreasing in θ. Further, if $\sigma = 0$ (or $\epsilon = 1$, as in the Cobb-Douglas case), the equilibrium becomes independent on the sectoral distribution of productivity and the degree of IPRs protection.

The long-run free trade equilibrium can now be found in Figure 2.1 as the intersection of the two schedules Φ (2.19) and TB (2.22). The graph can be used to study the effects of a strengthening of IPRs in the South. From (2.22), this implies a downward shift of the TB schedules which raises the relative wage in the South and reduces the set of goods produced there (z increases). Vice versa, a reduction of θ leads to a deterioration of the Southern relative wage and a relocation of some industries from the North to the South. Comparing (2.22) with (2.15), and noting that $\lim_{\theta\to 0}\omega = \max \phi_N(i)/\phi_S(i)$, proves the following:

Proposition 1 *For any $\sigma \in (0,1)$, there exists a level $\overline{\theta}$ such that if $\theta < \overline{\theta}$ income differences under free trade, as measured by ω, are larger than income differences under autarky.*

This is the first result of the paper, that trade can lead to divergence in income and productivity levels. Proposition 1 is based on the interplay between specialization and weak IPRs in developing countries: first, trade and specialization imply that the North and South benefit directly from different sets of innovations. Second, weak IPRs make innovations directed to the South less profitable. As $\theta \to 0$, R&D is directed towards Northern sectors only and the income gap grows up to its maximum $(\phi_N(0)/\phi_S(0))$, irrespective of any other country characteristics. In autarky, instead, even with $\theta = 0$, the South benefits from the innovation activities performed in all the sectors for the Northern market.

If North-South trade (with a low θ) shifts technology systematically in favor of the North, is it always beneficial for advanced countries? The striking answer is negative, as divergence opens the door to stagnation. To see this, calculate the equilibrium growth rate in free trade (see the Appendix for the derivation):

$$g^{FT} = L_N \left[\int_0^z \phi_N(i)^{\frac{\sigma}{(1-\sigma)}}\, di\right]^{\frac{1-\sigma}{\sigma}} \left(1 + \frac{L_S}{L_N}\frac{1}{\omega}\right)^{1/\sigma} - \rho. \tag{2.23}$$

Note that the growth rate of the world economy is increasing in θ: a higher θ expands the range z of goods produced in the North and decreases ω, all effects that contribute to raising the growth rate in (2.23). The intuition is simple and is the common argument in favor of IPRs protection: better enforcement of IPRs strengthens the incentives to innovate and therefore fosters growth. *But the surprising implication of (2.23) is that the growth rate of the world economy approaches* *zero if θ is low enough.* Endogenous growth is here possible because both the North and the South are growing. If innovations were not directed to Southern sectors, the Northern economy would be trapped into decreasing returns, not only because its sectors would experience falling output prices and profit margins, but also because more and more sectors would move to the South, where production is increasingly cheaper. In fact, long-run growth can stop even if $\theta > 0$. To see this, note that along the balanced growth path innovation has to be equally profitable in all the sectors; if θ is low enough, profitability of R&D in the South becomes so low that returns from investment fall short of the discount factor ρ and growth is destined to cease. Note that this result, like Proposition 1, requires $\sigma > 0$ (i.e., an elasticity of substitution between goods larger than one): with $\sigma = 0$ the cut-off commodity z and the wage ratio ω would not depend on technology, because every country and sector would benefit equally from any improvement in $a(i)$, and (2.23) would not depend on θ. Also, sector-specific technical process is a key assumption for deriving Proposition 2. In a set-up with factor-specific innovations, as in Acemoglu and Zilibotti (2001), the market size for any innovation depends on endowments only that are unaffected by specialization and trade: for this reason, incentives to invest in R&D would never go to zero even if $\theta = 0$.[21]

Comparing the growth rate in free trade, (2.23), and autarky, (2.14), and noting that (2.23) is a continuous function of θ with $\lim_{\theta \to \theta^* > 0} g^{FT} = 0$, proves the following:

Proposition 2 *For any $\sigma \in (0, 1)$, there exists a level $\widehat{\theta}$ such that, if $\theta < \widehat{\theta}$, the world growth rate is lower under free trade than under autarky.*

What happens during the transitional dynamics from autarky to the free trade equilibrium? Since technology adjusts slowly, initially the equilibrium is determined by equations (2.19) and (2.20) using the pre-trade values of $a(i)$. In general, the wage in both countries will jump up, as specialization increases the overall efficiency of the whole economy. Then, if the instantaneous wage ratio falls short of its long run free-trade value, there will be a period in which innovation is biased towards Northern sectors. During the transition, the Northern relative wage will rise and at

[21] As a consequence, in Acemolgu and Zilibotti (2001) trade opening has no effect on the world growth rate.

the same time firms will move to the South where production costs are lower. Note that in a trading environment with asymmetric IPRs protection, divergence and stagnation are closely related: it is the growing cost of producing in the wealthier North that induces the relocation of production towards the South (an important phenomenon in recent years) which in turn makes more sectors subject to weak IPRs and lowers the global incentives for innovation.

2.3 Why Are IPRs Not Protected in the South?

The previous analysis suggests that Southern countries may benefit from the enforcement of IPRs: it would attract more appropriate innovations and foster world growth. It is then interesting to ask why these policies are often not adopted. A first reason is that imitating countries would lose some profits: a marginal increase in θ induces a profit loss of $\beta\,(1-\beta)\,Y_S d\theta$, thereby reducing a country consumption level. Therefore, it can be optimal from the point of view of the South not to have full protection of IPRs. This is more likely the higher the profit share in the economy. Even if strong protection of IPRs is in the interest of the South, in the sense that the productivity gain due to higher or more appropriate innovation outweights the profit loss, the government might fail to implement the optimal policy for political reasons: if the group of monopolists that enjoy the rents from imitation has more political power that the workers, it may prefer to defend its share of profits at the expenses of the rest of the economy. Further, if the Southern policy makers behave myopically and fail to consider the effect of their policies on world innovation, then they would set an inefficiently low level of IPRs protection. Finally, in implementing IPRs protection, there might be a coordination problem among Southern governments of similar countries: each of them prefers the others to enforce IPRs, in order to attract innovation, but has an incentive to free ride not enforcing these property rights itself. However, this depends on the pattern of specialization and on the size of each country. If each Southern country specialized in a different set of commodities, then the coordination problem would disappear, as stronger IPRs would be beneficial for the enforcing country only. Similarly, a large country would have a higher incentive to protect IPRs because of its larger impact on world innovation and its limited ability to benefit from others' policies. To better understand these implications, the analysis is now extended to a multi-country setting.

2.4 Extensions

This section provides a sketch of how to extend the results to a multi-country world and how to incorporate trade barriers. These extensions add more realistic features to the basic model and help to clarify some of its empirical predictions. Consider first a case where the world economy can be divided into three homogenous regions: high (H), middle (M) and low (L) income countries. A key assumption here is that countries belonging to different regions have different exogenous productivities. The autarky solution is straightforward. To keep the analysis under free trade as simple as possible, assume that $\phi_H(i)/\phi_M(i)$ and $\phi_M(i)/\phi_L(i)$ are continuous and strictly decreasing in i. Further, assume that $\phi_H(i) > \phi_M(i) > \phi_L(i)$, $\forall i \in [0,1]$, implying that $w_H > w_M > w_L$ and that region H specializes in the lower range of goods $[0, z_1]$, region M in an intermediate range $[z_1, z_2]$ and region L produces the high-index goods $[z_2, 0]$. In this case, the first condition for a trading equilibrium, defining the cut-off sectors where it becomes profitable to move production form one region to another as a function of wages, becomes:

$$\frac{w_H}{w_M} = \frac{\phi_H(z_1)}{\phi_M(z_1)} \quad \text{and} \quad \frac{w_M}{w_L} = \frac{\phi_M(z_2)}{\phi_L(z_2)}.$$

The second equilibrium condition, trade balance, can be written in two equations:

$$w_H L_H \int_{z_1}^{1} p(i)^{1-\epsilon}\, di = w_M L_M \int_0^{z_1} p(i)^{1-\epsilon}\, di + w_L L_L \int_0^{z_1} p(i)^{1-\epsilon}\, di,$$

requiring the value of total imports in region H to be equal to the value of total export from region H; similarly for region L:

$$w_L L_L \int_0^{z_2} p(i)^{1-\epsilon}\, di = w_H L_H \int_{z_2}^{1} p(i)^{1-\epsilon}\, di + w_M L_M \int_{z_2}^{1} p(i)^{1-\epsilon}\, di.$$

Trade balance in region M is redundant. For a given technology and using (2.7) to substitute prices away, this system of four equations in four unknown (w_H/w_M, w_M/w_L, z_1 and z_2) can be solved to find the static equilibrium. Along the balanced growth path, innovation has to be equally profitable in all the sectors. In particular, considering sectors localized in different regions, the following condition must hold:

$$\theta_H \frac{\partial \pi_H(i)}{\partial a(i)} = \theta_M \frac{\partial \pi_M(j)}{\partial a(j)} = \theta_L \frac{\partial \pi_L(v)}{\partial a(v)},$$

for any i, j, v such that $i \leq z_1 \leq j \leq z_2 \leq v$. These conditions can be used to characterize the new trading equilibrium. Leaving the details of the analysis aside, it is easy to see how the logic of previous results extends to the multi-country setting: because of specialization, under free trade a tightening of IPRs in a region (or in a large country of the region) attracts more innovation towards the goods the region is producing. This translates into a higher wage and a reduction of the range of activities performed in the region (moving production abroad becomes more convenient as the domestic labor cost increases). On the contrary, the positive effects of tighter IPRs in a region in autarky are spread across all sectors and affects only a small fraction of the market for innovations (the fraction of profits coming from that specific region) and therefore are less likely to have a significant impact on world incentives to innovate. The main result of the basic model is therefore reinforced: because of specialization, country regulations become more effective in an integrated economy.

Finally, it is easy to see how the introduction of trade costs gives rise to a range of non-traded goods and hence a regime that combines elements of both the free-trade and autarky equilibrium. Assume that trade costs take the following "iceberg" form: of a unit of good shipped from a country, only a fraction $1/\tau$ ($\tau \geq 1$) arrives at destination. Consider then the conditions for an efficient allocation of production in the simple two-country world. The North will export a good only if it is competitive in the foreign market after taking into account the trade cost. Therefore, the borderline commodity where exports stop must now satisfy the following condition:

$$\omega = \frac{\phi_N(\underline{z})}{\tau \phi_S(\underline{z})} \tag{2.24}$$

i.e., comparative advantage, reduced by the trade cost, has to be equal to the competitiveness of the North (the relative production costs, ω). Similarly, a good will be imported if the foreign price, multiplied by τ, is lower than the domestic production cost. Again, the threshold commodity is identified by the indifference between home and foreign production:

$$\omega = \frac{\tau \phi_N(\overline{z})}{\phi_S(\overline{z})} \tag{2.25}$$

For a given relative wage ω, equations (2.24) and (2.25) divide the commodity space into three categories: goods $[0, \underline{z}]$ produced in the North only, goods $[\overline{z}, 1]$ produced in the South only and a range $[\underline{z}, \overline{z}]$ of non-traded goods produced in both regions for the domestic markets. Trade balance is now referred to traded goods only:

$$w_N L_N \int_{\overline{z}}^{1} p(i)^{1-\epsilon} \, di = w_S L_S \int_{0}^{\underline{z}} p(i)^{1-\epsilon} \, di$$

Finally, along a balanced growth path, innovation is performed in all the sectors. Considering sectors in all the possible sets of goods, the arbitrage condition for R&D becomes:

$$\theta_N \frac{\partial \pi_N(i)}{\partial a(i)} = \theta_N \frac{\partial \pi_N(j)}{\partial a(j)} + \theta_S \frac{\partial \pi_S(j)}{\partial a(j)} = \theta_S \frac{\partial \pi_S(v)}{\partial a(v)}$$

for $i \in [0, \underline{z}]$, $j \in [\underline{z}, \overline{z}]$ and $v \in [\overline{z}, 0]$. Note that as trade costs vanish, $\tau \to 1$, the range of non traded goods disappears ($\underline{z} \to \overline{z}$) and economy approaches the free trade equilibrium. Conversely, as barriers become prohibitive, $\tau \to \infty$, all the goods turn non traded ($\underline{z} \to 0$ and $\overline{z} \to 1$) and the economy converges to the autarky case.

3 Empirical analysis

The key mechanism of the model is the interaction between trade-driven specialization and the ability of a country to attract better technologies by changing the level of protection of IPRs. Given an elasticity of substitution across sectors larger than one ($\epsilon > 1$ or $\sigma > 0$), more innovation targeted to a sector translates into higher sectoral income, both in absolute terms and relative to the rest of the economy. Because of this, a country unambiguously gains from innovations on the goods it is producing. Innovation, in turn, can be stimulated by protecting more the rewards of inventors. In this set-up, specialization has two effects. First, by increasing a country's share of world production (and profits) in the sectors of specialization, it increases the impact of country policies on global profitability of innovations directed to those sectors, thereby increasing the ability of a country to attract technologies tailored to its needs. Second, by reducing the number of countries producing a specific good, it limits the benefits of innovations directed to that good on the rest of the world. For these reasons, the model suggests the positive effect of raising θ_i on income of country i to be higher under free trade than in autarky or, more generally, the larger the range T of traded goods in the economy. Further, since the ability of country i to attract innovation in sector j depends on its share in world production of that sector, which in turn depends on country size, the model suggests that the impact of θ_i on productivity should be higher in larger countries. These implications can be summarized as follows:

$$\frac{\partial \left(Y_i/L_i\right)}{\partial\theta_i\partial T} > 0 \quad \text{and} \quad \frac{\partial \left(Y_i/L_i\right)}{\partial\theta_i\partial L_i} > 0 \tag{2.26}$$

Since the main results hinge critically on these interactions, testing the sign of the cross-partial derivatives in (2.26) provides a way to assess the empirical plausibility of the model. Predictions on the overall effect of IPRs seem instead less useful to evaluate the theory. Although the model implies that raising θ should always have a positive effect on productivity, this result relies heavily on the simplifying assumption that θ does not affect the monopoly distortion in the South.

To test (2.26), measures of labor productivity, IPRs protection, openness to trade and size have been collected for a panel of countries from 1965 to 1995. Labor productivity is proxied by real GDP per worker (GPDW) from the Penn World Table 6.0 (PWT6.0). Two important determinants of productivity are also included in the analysis: the stock of physical capital per worker (KL) again from PWT6.0 and the fraction of working age population with at least secondary schooling as a proxy for human capital (HL), from Barro-Lee. As for trade openness, two different measures are considered: the Sachs and Warner (1995) index, which is a dummy taking value one if a country is classified as open,[22] and the trade share in total GDP form PWT6.0. Although the first is useful to distinguish countries under different trade regimes, it has almost no time variation and is therefore appropriate to explain

[22]An economy is classified as open if satisfies all of the following criteria: (1) nontariff barriers cover less than 40 percent of trade (2) average tariff rates are less than 40 percent (3) any black market premium was less than 20 percent during the 1970s and 1980s (4) the country is not classified as socialist and (5) the government does not monopolize major exports.

only the cross-section. The second measure, instead, captures well the increase in market integration over time. Country size is measured by total population (POP), as reported in PWT6.0. The last challenge is to find reliable data on the degree of protection of intellectual property. In this respect, this study uses the index of patent rights built by Ginarte and Park (1995). Although patents are only a component of IPRs, they are likely to be highly correlated with the overall level of protection; further, this index has the advantages of being available for a large number of countries with quinquennial observation since 1965 and of being based on both the strength and enforceability of national laws.[23] The index (IPR) ranges from 0 to 5. In summary, the overall dataset comprises a cross-section of 53 countries and 6 time observations, from 1965 to 1990 at 5 year intervals.[24] Descriptive statistics are reported in Table 1.

[23]This index is based on an assessment of five aspects of patent laws: (1) extent of coverage, (2) membership in international patent agreements, (3) provision for loss of protection, (4) enforcement mechanisms and (5) duration of protection. An alternative, but time-invariant, measure of IPRs is provided by Rapp and Rozek (1990). On the cross-section, the two proxies yield very similar results.

[24]Data are available for the following countries: Argentina, Australia, Austria, Belgium, Bolivia, Botswana, Canada, Chile, Colombia, Denmark, Dominican Rep., Ecuador, Finland, France, Greece, Guatemala, Honduras, Hong Kong, Iceland*, India, Iran, Ireland, Israel, Italy, Jamaica, Japan, Kenya, Korea Rep., Malawi, Mauritius, Mexico, Nepal, Netherlands, New Zealand, Norway, Panama*, Paraguay, Peru, Philippines, Portugal, Sierra Leone, Spain, Sri Lanka, Sweden, Switzerland, Syria, Thailand, Turkey, U.K., U.S.A., Venezuela, Zambia, Zimbabwe. An asterisk (*) indicates no Sachs and Warner index available.

Table 1: Descriptive Statistics

	IPR	OPEN*	OPEN	KL	HL	POP	GDPW
1965	2.47 (0.59)	0.52 (0.50)	46.69 (25.69)	7848 (7703)	19.82 (18.39)	26420 (70771)	16953 (11608)
1970	2.52 (0.67)	0.51 (0.50)	50.37 (29.52)	10232 (9265)	23.51 (19.61)	29003 (78764)	18915 (12248)
1975	2.53 (0.67)	0.49 (0.50)	57.83 (29.51)	12997 (11394)	26.11 (19.95)	31833 (87549)	20917 (13244)
1980	2.69 (0.85)	0.52 (0.50)	61.42 (31.38)	15190 (12781)	32.72 (22.09)	34782 (97354)	21347 (14101)
1985	2.71 (0.89)	0.49 (0.50)	60.69 (35.42)	16507 (14154)	35.59 (21.63)	37821 (107662)	23412 (15666)
1990	2.75 (0.90)	0.70 (0.46)	63.54 (38.14)	18754 (16336)	40.26 (21.99)	41039 (118867)	25433 (16960)
Correlation Matrix							
IPR	1.00						
OPEN*	0.40	1.00					
OPEN	0.20	0.26	1.00				
KL	0.55	0.50	0.11	1.00			
HL	0.61	0.50	0.16	0.78	1.00		
POP	-0.05	-0.07	-0.31	-0.07	-0.01	1.00	
GDPW	0.59	0.60	0.16	0.86	0.80	-0.05	1.00

Note: OPEN* is the Sachs and Warner index of openness. Standard error in parentheses.

To get a first sense for the patterns in the data, Table 2 presents a set of conditional correlations. The results are encouraging for the present theory. As predicted by the model, IPRs protection is associated with higher productivity only for countries classified as open by Sachs and Warner. The correlation is zero for closed economies. Likewise, being open has a much higher correlation with productivity in countries with strong patent rights. Also the second prediction in (2.26) seems broadly consistent with the data, as IPRs protection is found to have a higher correlation with productivity in larger countries.

Table 2: Conditional Correlations

Variable	Conditional on	CORR with GDPW	N. obs.
IPR	OPEN=0	0.003	146
IPR	OPEN=1	0.748	166
OPEN	IPR<2.5	0.238	135
OPEN	IPR>=2.5	0.726	177
IPR	POP<mean	0.48	254
IPR	POP>=mean	0.85	70

Note: OPEN= Sachs and Warner index of openness

A better way to display these correlations is through simple least-square regressions on the pooled data. Throughout, all the variables are in logs, except for dummies; further, to alleviate simultaneity concerns, all the right-hand side variables are lagged five years. Column (1) of Table 3 reports the results of regressing real output per worker (GDPW) on patent rights (IPR) the Sachs and Warner openness index (OPEN), an interaction term between IPR and OPEN, an interaction term between IPR and country size (POP) and country size itself (POP). The regression also controls for the two important determinants of productivity, physical (KL) and human (HL) capital per worker. According to (2.26) the two interaction terms should have a positive sign. Consistently, column (1) shows that the coefficient on both interactions is positive and precisely estimated.

Although the pooled OLS regression is a useful way to summarize partial correlations in the data, it may place too much weight on cross-sectional variation and suffer from omitted variables, particularly given the small number of covariates. In this respect, a LSDV regression with country fixed-effects has more advantages, as it controls for omitted variables that change very little over time and that may be correlated with other regressors, such as institutional and geographical characteristics of countries. However, since this estimator uses only within-country variation, the Sachs and Warner index of openness, with its almost nil time variation, is here inadequate. The analysis therefore continues using the trade share in GDP as a measure of openness. Before moving to the fixed-effects regression, Column (2) shows again the pooled OLS estimates with the new trade measure and it confirms the previous findings: the two interaction terms are positive and significant at the 1% level.

Table 3: Panel Analysis

	OLS(1)	OLS(2)	LSDV(3)	LSDV(4)	LSDV(5)
IPR	-1.941 (0.697)***	-5.723 (1.568)***	-0.407 (0.875)	-0.464 (0.411)	-0.904 (0.488)*
OPEN	-0.437 (0.200)**	-0.719 (0.231)***	0.041 (0.098)	0.038 (0.096)	0.153 (0.114)
IPR*OPEN	0.801 (0.265)***	0.556 (0.212)***	0.216 (0.105)**	0.219 (0.103)**	0.385 (0.122)***
IPR*POP	0.163 (0.065)**	0.393 (0.089)***	-0.005 (0.074)	-	-
POP	0.207 (0.70)***	-0.452 (0.092)***	-0.013 (0.113)	-	-
KL	0.400 (0.075)***	0.453 (0.073)***	0.323 (0.034)***	0.321 (0.031)***	-
HL	0.164 (0.084)*	0.214 (0.080)***	-0.037 (0.036)	-0.042 (0.024)*	-
R^2	0.83	0.82	0.58	0.58	0.39
No. of Obs.	306	318	318	318	318
F-test[country effects] (P-value)	-	-	31.02 (0.000)	39.06 (0.000)	122.44 (0.000)
Hausman χ^2 (P-value)	-	-	176.16 (0.000)	48.56 (0.000)	8.63 (0.034)

LHS: real GDPW. All variables, except dummies, in logs. RHS variables are lagged (5 yeras). Column 1 uses the Sachs and Warner Openness index. Columns 2-5, use the trade share in GDP. Standard errors in parenthesis (robust, in OLS regressions). Constant not reported. *, ** and *** indicate significance at 10%, 5% and 1% level.

Columns (3)-(5) report the results from the LSDV fixed-effects estimator. Column (3) includes all the right-hand side variables. The interaction term between patent rights and openness is still positive and significant. On the contrary, the coefficient on country size is now very small and not statistically different from zero. This is not very surprising, given that population varies mostly across countries (Table 1 shows that the cross-sectional standard error of POP is almost three times its mean). It suggests that only the large cross-sectional variation of country size may have a significant impact on the effectiveness of IPRs, which is not inconsistent with the theory. Column (4) reports the estimates after dropping the size variables, whose contribution to explain changes in productivity over time has been found statistically small. Finally, Column (5) isolates the effects of patent rights and trade, the main variables of interest, by dropping all the other covariates. In all cases, the coefficient on the interaction term between openness and patent rights is consistently found to be positive and statistically different form zero.[25] To conclude, given that in all the specifications the coefficient on the interaction term is found to be positive with significance levels always below 4%, there seems to be fairly robust evidence that patent laws are more correlated with high productivity in open countries.

A few calculations on the coefficients in Table 3 can help to understand the magnitude of the effects and if the estimates across specifications are comparable. Consider first the impact of intellectual property protection. For the average country, Columns 1-3 imply that a 10% increase of the index of patent rights is associated with an output change of -0,3%, +0,7% and +3,8% respectively. These numbers suggest that, for the average country, gains form stronger IPRs may be uncertain. The situation is different for trading economies: with openness one standard error above the sample mean, the reaction of output becomes +3,7%, +4% and +5,1% respectively. Conversely, for countries closed to trade (one standard deviation below the sample mean) the effect may be negative: -4,3%, -2,5% and +2,5%. Similarly, according to Columns 1-3, a 10% increase of the openness index in the average country is associated with an output change of +2,9%, -2,1% and +1,5%, respectively. In countries with patent rights one standard error above the sample mean, the positive effect of trade is instead more pronounced: +5,5%, -0,3% and +2,2%. Finally, for countries with patent rights one standard error below the sample mean, the effect of trade becomes small or even negative: +0,3%, -3,9% and +0,8%. Although the variability of estimates across specifications is not too high, given that coefficients come form regressions using very different trade measures and estimation techniques, it makes it difficult to draw sharp empirical conclusions. However, these numbers indicate that open and perhaps large economies may benefit substantially from stronger patent laws. It may thus suggest that the process of trade liberalizations in India and China could be more beneficial if accompanied by a tightening of IPRs. Moreover, given the 34% increase of average openness over the sample

[25] Adding a time trend affects the results only marginally and turns out not significant.

period and the high correlation between patent rights and income, these estimates suggest that globalization, coupled with low protection of intellectual property in poor countries, may have contributed to the widening of living standards.

How do these results relate to the empirical literature on trade, growth and convergence? A general finding of several influential papers is that openness promotes growth and convergence. In particular, a first strand of literature documents a positive correlation between trade and growth.[26] Likewise, this paper shows that integration may enhance productivity in all countries because of static (and potentially dynamic) gains from trade, but in addition it argues that countries with better IPRs policies may reap more benefits than others. Further, recent works by Easterly and Levine (2002) and Rodrik et al. (2002) have questioned the robustness of the correlation between trade and growth. In particular, these authors argue that the correlation disappears after controlling for institutional quality and addressing endogeneity issues. The importance of institutions is again in line with the central message of this paper: that the effect of trade on productivity and growth depends crucially on property rights, which are an important institutional factor. A second strand of literature is focused on market integration and convergence. Here, Barro and Sala-i-Martin (1995) find strong evidence of convergence among highly integrated countries and regions (OECD countries, the US states, European regions and Japanese prefectures) and Ben-David (1993) shows the removal of trade barriers fostered convergence across countries who joined the European Economic Community. These results are not inconsistent with the model and the evidence presented in this paper, because they show the pro-convergence effect of integration between countries with similar property rights related regulations.

Before concluding, it is worthwhile to mention briefly some interesting empirical observations. The model predicts that in a period of growing world trade the R&D effort of advanced countries should become more specialized towards the sectors in which those countries have a comparative advantage. In this respect, it is perhaps suggestive to look at the evolution of the number of patents by technological category issued in the US over the last four decades, reported by Hall et al. (2001): the three traditional fields (Chemical, Mechanical and Others) have experienced a steady

[26] Frankel and Romer (1999) and Sachs and Warner (1995) are two notable examples.

decline, dropping from a share of 76% of total patents in 1965 to 51% only in 1990. Conversely, Computers and Communications rose from 5% to over 20%, Drugs and Medical form 2% to 10%, whereas Electrical and Electronics is the only stable field (16-18% of total). Albeit consistent with the theory, this evidence is more difficult to interpret, as it may reflect technology cycles or changes in demand. More in general, the model generates something resembling a product cycles, where sectors become less technology intensive *after* they move to the South. Distinguishing empirically between this prediction and the traditional view, according to which goods become less technology intensive *before* moving to LDCs, seem an interesting challenge for future work.

4 Concluding Remarks

This paper has presented a simple model where market integration can amplify income differences between rich and poor countries and lower the world growth rate, even in the presence of standard mutual gains from trade. Rather than raising warnings against globalization, the analysis has identified a specific market failure, weak protection of intellectual property in developing countries, under which trade can have undesirable effects. In a world of integrated economies, profits from innovations play a crucial role in directing technical progress towards the needs of all countries and in sustaining long-run growth incentives. This suggests that trade liberalization in developing countries should be accompanied by reforms aimed at a tightening of intellectual property rights. With the inclusion of the TRIPS agreement in the WTO, international negotiations have recently taken important steps in that direction. A major contribution of this paper was thus to provide new theoretical foundations for these efforts. However, even though the analysis hints at large potential gains from global regulations, imposing common standards can be costly for some less developed countries and may not be sufficient. As long as the economic weight of the South is low, profits generated from its markets would not be enough to provide the right incentives for developing appropriate technologies. Although the model has focused on intellectual property rights asymmetries, the sale of innovations in poor countries can generate small profits for a number of other reasons, including high transaction costs and risks of expropriation. Given these distortions, promoting research aimed at the needs of the less developed countries appears to be a the key element for reducing cross-country income differences and fostering world growth.

While the paper has emphasized the quasi public good nature of technology emerging from the endogenous growth literature, where knowledge flows with no frictions across borders, trade itself could contribute to technology transfer between countries. Similarly, the paper has abstracted from new products and product cycle trade. Further, infringements of intellectual property rights has been modeled in a very stylized way that does not explicitly include micro details. Incorporating these elements into the analysis would certainly help to understand the complex interactions between innovation and income in the global economy and seems a fruitful direction for future research. Finally, the paper has shown that the consequences of globalization may depend on institutional variables such as property right laws. Whether the effects described can be important in shaping the world income distribution and affecting innovating incentives, remains an empirical question that deserves further study.

5 References

Acemoglu, Daron, Philippe Aghion and Fabrizio Zilibotti (2002). "Distance to Frontier, Selection and Economic Growth," MIT, mimeo.

Acemoglu, Daron and Jaume Ventura (2002). "The World Income Distribution" *Quarterly Journal of Economics* 117, 659-694.

Acemoglu, Daron and Fabrizio Zilibotti (2001). "Productivity Differences," *Quarterly Journal of Economics* 116, 563-606.

Antras, Pol (2002). "Incomplete Contracts and the Product Cycle," MIT, mimeo

Barro, Robert J. and Xavier Sala-i-Martin (1995). *Economic Growth* McGraw-Hill.

Beaudry, Paul, Fabrice Collard and David A. Green (2002). "Decomposing the

Twin-peaks in the World Distribution of Output-per-worker" mimeo, University of British Columbia

Ben-David, Ben (1993). "Equalizing Exchange: Trade Liberalization and Income Convergence," *Quarterly Journal of Economics* 108, 653-679.

Bhagwati, Jagdish (1958). "Immiserizing Growth: a Geometrical Note," *Review of Economic Studies* 25, 201-205.

Chin, Judith C. and Gene M. Grossman (1990) "Intellectual Property Rights and North-South Trade" in *The Political Economy of International Trade: Essays in Honor of Robert E. Baldwin*, ed. by R.W. Jones and A.O. Krueger. Cambridge, MA, Basil Blackwell.

Deardorff, Alan V. (1992). "Welfare Effects of Global Patent Protection" (1992). *Economica* 59, 35-52.

Dinopoulos, Elias and Paul Segerstrom (2003). "A Theory of Globalization," Stockholm School of Economics, mimeo

Diwan, Ishac and Dani Rodrik (1991). "Patents, Appropriate Technology, and North-South Trade," *Journal of International Economics* 30, 27-48.

Dornbusch, Rudi, Stanley Fischer and Paul A. Samuelson, (1977) "A Continuum Ricardian Model of Comparative Advantage, Trade and Payments" *American Economic Review* 67, 823-839.

Easterly, William and Ross Levine (2002). "Tropics, Germs and Crops: How Endowment Influence Economic Development," mimeo, Center for Global Development and Institute for International Economics, 2002.

Frankel, Jeffrey and David Romer (1999). "Does Trade Cause Growth?" *American Economic Review* 89, 379-399

Gancia, Gino A. and Fabrizio Zilibotti (2003) "Endogenous Growth with Product Variety: Theory and Applications" mimeo

Ginarte, Juan C. and Walter G. Park (1997). "Determinants of Patent Rights: A Cross-National Study," *Research Policy* 26, 283-301.

Gould, M. David and William C. Gruben (1996). "The Role of Intellectual Property Rights in Economic Growth," *Journal of Development Economics* 48, 323-350.

Grossman, Gene M. and Edwin L. C. Lai (2002). "International Protection of Intellectual Property," NBER W.P. 8704.

Hall, Bronwyn H., Adam B. Jaffe and Manuel Trajtenberg (2001). "The NBER Patent Citations Data File: Lessons, insights and Methodological Tools," NBER WP8498.

Helpman, Elhanan (1993). "Innovation, Imitation and Intellectual Property Rights," *Econometrica* 61, 1247-1280

Jaffe, Adam B., Manuel Trajtenberg and Rebecca Henderson (1993). "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations," *Quarterly Journal of Economics* 108 577-598.

Kremer, Michael (2002). "Pharmaceuticals and the Developing World," *Journal of Economic Perspective* 16, 67-90.

Krugman, Paul R. (1979) "A Model of Innovation, Technology Transfer and the World Distribution of Income" *Journal of Political Economy* 87, 253-266.

Lai, Edwin L. C. (1998). "International Intellectual Property Rights Protection and the Rate of Product Innovation," *Journal of Development Economics* 55, 115-130.

Maddison, Angus (1995). *Monitoring the World Economy, 1820-1992.* Paris: Development Center of the Organization for Co-operation Development.

Maskus, Keith E. and Mohan Penubarti (1995). "How trade-related are intellectual property rights?" *Journal of International Economics* 39, 227-248.

Matsuyama, Kiminori (2000). "A Ricardian Model with a Continuum of Goods under Nonhomothetic Prefernces: Demand Complementarities, Income Distribution, and North-South Trade," *Journal of Political Economy* 108, 1093-1120.

Pritchett, Lant (1997). "Divergence Big Time" *Journal of Economic Perspective* 11, 3-17.

Quah, Danny (1993). "Empirical Cross-Section Dynamics in Economic Growth," *European Economic Review*, 37, 426-434.

Rapp, Richard T. and Richard P. Rozek (1990)."Benefits and Costs of Intellectual Property Protection in Developing Countries" *Journal of World Trade* 24, 74-102.

Rodrik, Dani, Arvind Subramanian and Francesco Trebbi (2002). "Institutions Rule: The Primacy of Institutions over Geography and Integration in Economic Development" mimeo, Harvard University.

Romer, Paul M. (1990). "Endogenous Technological Change," *Journal of Political Economy* 98, 71-102.

Sachs, Jeffrey D. (1999) "Helping the World's Poorest," *The Economist* 14 August

Sachs, Jeffrey D. and Andrew Warner (1995). "Economic Reform and the Process of Global Integration," *Brookings Papers on Economic Activity* 1995, 1-95.

Sala-i-Martin, Xavier (2002). "The Disturbing "Rise" of Global Income Inequality," mimeo Columbia University.

Taylor Scott M. (1994) "TRIPs Trade and Growth," *International Economic Review* 35, 361-381.

Wood, Adrian (1998). "Globalisation and the Rise in Labour Market Inequalities," *The Economic Journal* 108, 1463-1482.

Yang, Guifang and Keith E. Maskus (2001). "Intellectual property rights, licensing and innovation in an endogenous product-cycle model" *Journal of International Economics* 53, 169-187.

Young, Alwyn (1991) "Learning by Doing and the Dynamic Effects of International Trade," *Quarterly Journal of Economics* 101, 396-405.

6 Appendix

6.1 Optimality of technologies

Consider first the case of no IPRs protection in S, $(\theta = 0)$. Total production in the North is equal to $Y_N = w_N L_N/\beta$. Using (2.9):

$$Max_{\{a(i)\}} Y_N = L_N \left\{ \int_0^1 \left[a\left(i\right) \phi_N(i) \right]^\sigma di \right\}^{1/\sigma} \quad s.t. \quad \int_0^1 a\left(i\right) di = a$$

The solution to this program has to satisfy the following first order conditions (FOCs), $\forall i \in [0,1]$:

$$L_N \left\{ \int_0^1 \left[a\left(i\right) \phi_N(i) \right]^\sigma di \right\}^{\frac{1-\sigma}{\sigma}} \left[a\left(i\right) \phi_N(i) \right]^{\sigma-1} \phi_N(i) = \lambda$$

where λ is the lagrange multiplier associated to the constraint. Taking the ratio of any two FOCs and using $A_N(i) = a\left(i\right)\phi_N(i)$ yields equation (2.13). This proves that the sectoral profile of the endogenous technology maximizes Northern output and wage and hence it is optimal for the North.

Consider now the case of imperfect protection of IPRs in S, $(\theta \neq 0)$.

$$\begin{aligned} Max_{\{a(i)\}} Y_N + \theta Y_S &= L_N \left\{ \int_0^1 \left[a\left(i\right) \phi_N(i) \right]^\sigma di \right\}^{1/\sigma} + \theta L_S \left\{ \int_0^1 \left[a\left(i\right) \phi_S(i) \right]^\sigma di \right\}^{1/\sigma} \\ s.t. \quad \int_0^1 a_N\left(i\right) di &= a \end{aligned}$$

the FOCs for a maximum are, $\forall i \in [0,1]$:

$$\begin{gathered} L_N \left\{ \int_0^1 \left[a\left(i\right) \phi_N(i) \right]^\sigma di \right\}^{\frac{1-\sigma}{\sigma}} \left[a\left(i\right) \phi_N(i) \right]^{\sigma-1} \phi_N(i) + \\ \theta L_S \left\{ \int_0^1 \left[a\left(i\right) \phi_S(i) \right]^\sigma di \right\}^{\frac{1-\sigma}{\sigma}} \left[a\left(i\right) \phi_S(i) \right]^{\sigma-1} \phi_S(i) = \lambda \end{gathered}$$

where λ is the lagrange multiplier associated to the constraint. Using (2.9) and solving for $a\left(i\right)$:

$$a\left(i\right) = \left[\frac{L_N \phi_N\left(i\right)^\sigma \left(w_N\right)^{1-\sigma} + \theta L_S \phi_S\left(i\right)^\sigma \left(w_S\right)^{1-\sigma}}{\beta\lambda} \right]^{1/(1-\sigma)}$$

Comparing this condition with equation (2.16) in the text shows that the sectoral distribution of the endogenous technology maximizes a weighted sum of Northern and Southern aggregate output, with a weight of θ on the South. As $L_N/(\theta L_S) \to 0$, technologies maximize w_S, whereas as $L_N/(\theta L_S) \to \infty$ they maximize w_N.

6.2 Properties of the wage ratio in autarky

To show that the North-South wage ratio in autarky is bounded by $\max \phi_N(i)/\phi_S(i) = \phi_N(0)\,\phi_S(0)$, first note that $\partial\omega/\partial\phi_N(i) > 0$ and $\partial\omega/\partial\phi_S(i) < 0$. Therefore, by construction:

$$\omega = \left[\frac{\int_0^1 \phi_N(i)^{\sigma/(1-\sigma)}\, di}{\int_0^1 \phi_N(i)^{\sigma^2/(1-\sigma)}\, \phi_S(i)^{\sigma}\, di}\right]^{1/\sigma} \leq \left[\frac{\int_0^1 \phi_N(0)^{\sigma/(1-\sigma)} di}{\int_0^1 \phi_N(0)^{\sigma^2/(1-\sigma)}\, \phi_S(0)^{\sigma}\, di}\right]^{1/\sigma} = \frac{\phi_N(0)}{\phi_S(0)}$$

6.3 The growth rate under free-trade

Rewrite the marginal condition for buying innovation in a Northern sector as:

$$\frac{w_N \phi_N(i)\, L_N A_N(i)^{\sigma-1}}{\beta \int_0^z A_N(j)^{\sigma}\, dj} = r$$

use (2.7) to substitute for w_N. Rearrange it to get:

$$p(i)^{1-\epsilon} = \left[\frac{\phi_N(i)\, L_N A_N(i)^{\sigma}}{r \int_0^z A_N(j)^{\sigma}\, dj}\right]^{\sigma}$$

use $A_N(j) = A_N(i)\left[\frac{\phi_N(j)}{\phi_N(i)}\right]^{1/(1-\sigma)}$ to elimnate $A_N(i)$. Integrate i over the interval $[0,1]$, use (2.3) and rearrange:

$$r = \left\{(L_N)^{\sigma}\left[\int_0^z \phi_N(i)^{\frac{\sigma}{1-\sigma}}\, di\right]^{1-\sigma} + (\theta L_S)^{\sigma}\left[\int_z^1 \phi_S(i)^{\frac{\sigma}{1-\sigma}}\, di\right]^{1-\sigma}\right\}^{1/\sigma}$$

Finally, use (2.22) to substitute for $\int_z^1 \phi_S(i)^{\sigma/(1-\sigma)}\, di$. The Euler equation $g = r - \rho$ then yields equation (2.23) in the text.

Chapter 3

The Skill Bias of World Trade*

1 Introduction

Wage inequality has widened over the last two decades. This fact has stimulated a growing body of research, which has pointed at skill-biased technical change and international trade as major explanations. It has been argued that technology can be at the root of the increase in inequality because recent innovations in the production process, such as the widespread introduction of computers, have boosted the relative productivity of skilled workers.[1] In contrast, trade models generally attribute the rising skill premium in OECD countries to the growing competition with imports from low-wage producers due to globalization.[2] However, the current consensus is that the role of international trade has little empirical relevance compared to the role of technology. There are three main reasons why the conventional trade explanation fails to convince. First, although the last two decades have witnessed a substantial increase in the volume of North-South trade, advanced countries still

* Written with Paolo Epifani. We are grateful to Olivier Blanchard, Giovanni Bruno, Francesco Caselli, Alan Deardorff, Henrik Horn, Gianmarco Ottaviano, Bob Staiger, Alessandro Turrini, Dieter Urban, seminar participants at IIES, Stockholm University, MIT, the EEA Annual Meeting (Venice, 2002), the CNR Conference (Milano, 2001) and especially Daron Acemoglu, Torsten Persson, Jaume Ventura and Fabrizio Zilibotti for helpful comments. We thank Christina Loennblad for editorial assistance. The usual caveat applies. Financial support from the Wallander and Hedelius Foundation is gratefully acknowledged. Gino Gancia thanks the MIT Economics Department for hospitality.

[1] See, among others, Autor et al. [1998].

[2] In particular, Wood [1994, 1998] proposes an augmented Heckscher-Ohlin theory based on specialized trading equilibria. Feenstra and Hanson [1996, 1999] instead emphasize the role played by intensive outsourcing of less skill-intensive activities.

trade too little with developing countries for the effect of low-price imports to be quantitatively relevant.[3] Second, the rise in the skill premium has also occurred in many developing countries, which runs counter to the conventional trade story.[4] Third, most studies suggest that the relative price of skill-intensive goods did not increase during the period of rising inequality,[5] whereas trade models imply an unambiguous positive relationship between prices of factors and goods.

In this paper, we propose a new role of international trade in explaining wage inequality consistent with the empirical evidence. We do so by revisiting the *new trade theory*'s account of the distributional effects of intra-industry trade. By definition, intra-industry trade is trade in goods with similar factor intensities; therefore it has no impact on relative factor demand and cannot explain the evolution of the skill premium according to conventional wisdom. We argue that this seemingly plausible result hinges either on Cobb-Douglas preferences or perfect symmetry between sectors. We show that an elasticity of substitution in consumption greater than one and higher scale economies in the skill-intensive sectors imply that any increase in the volume of trade, even between identical countries, tends to be skill-biased. The intuition behind this result is very simple. Trade expands the market size of the economy, which is beneficial because of increasing returns. In relative terms, however, output increases by more in the skill-intensive sector, since it is characterized by stronger economies of scale, and the relative price of the skill-intensive good therefore falls. With an elasticity of substitution in consumption greater than one, the demand for skill-intensive goods increases more than proportionally, raising their share of total expenditure and therefore also the relative wage of skilled workers.

This result has important implications. First, it suggests that the entire volume of world trade matters for inequality and not only the small volumes of North-South trade. We show that the skill bias of trade is quantitatively relevant; under reason-

[3] Wood [1998] reports that imports of manufactures from developing countries constitute a small fraction of OECD GDP (about 3%), although this share has almost tripled between 1980 and 1995. The point that these volumes of trade are too small to have an important effect on wage inequality has been forcefully made by Krugman [2000]. See also Deardorff [2000] and Deardorff and Staiger [1988] on this point.

[4] For evidence on wage inequality in developing countries see Robbins [1996], Hanson and Harrison [1999] and Berman, Bound and Machin [1998].

[5] Lawrence and Slaughter [1993], in particular, document a decline in the relative price of US skill-intensive goods in the 1980s.

able parameter values trade integration between two identical countries can increase skill premia by almost 10%. Second, if the skill-biased scale effect is strong enough to overcome the standard factor proportions effect, international trade will spur inequality even in the skill-poor developing economies, making the model consistent with the evidence of rising skill-premia in developing countries that have experienced trade liberalizations. In particular, we show in a simple numeric exercise that trade integration between Mexico and the United States can account for a significant increase in the Mexican skill premium. Third, our model can explain the decline in the relative price of skill-intensive goods during the period of rising skill premia and growing volumes of world trade. In the framework we propose, the so-called price puzzle (the empirical finding that relative factor and good prices moved in opposite directions) simply disappears.

We also extend our analysis by introducing physical capital. As the capital stock is an important component of economic size, we find that its accumulation leads to higher skill premia. More interestingly, we show that the intersectoral mobility of capital is likely to magnify the effects of trade integration on wage inequality. Our findings are consistent with both the evidence on capital relocation towards skill-intensive sectors (Caselli, [1999]) and the large literature on capital-skill complementarity.

As mentioned, our results rest on scale economies being relatively stronger in the skill-intensive sectors and the elasticity of substitution between goods of different skill-intensity being greater than one. How realistic are these assumptions? Paul and Siegel [1999] estimate returns to scale in US manufacturing industries at the two-digit industry level for the period 1979-1989. Figure 3.1 plots their estimates against a measure of sectoral skill-intensity. For each industry, the vertical axis reports the output elasticity of the long-run total cost function (an inverse measure of internal and external scale economies) and the horizontal axis the share of production workers in total employment in 1990 (an inverse measure of skill-intensity). The diagram clearly shows a positive correlation between skill-intensity and scale economies. We also report a weighted regression line, whose slope coefficient and standard error are 0.59 and 0.21, respectively. Similar results are reported by Antweiler and Trefler [2002]; using international trade data for 71 countries and a very different method-

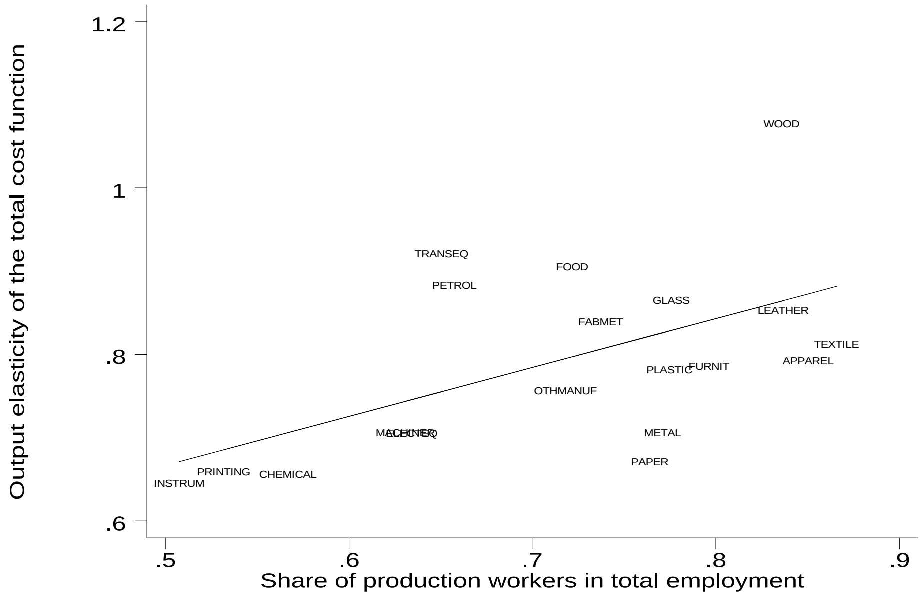

Figure 3.1: Skill-intensity and increasing returns

ology, they find that skill-intensive sectors, such as Petroleum Refineries and Coal Products, Pharmaceuticals, Electric and Electronic machinery and Non-Electrical Machinery, have an average scale elasticity around 1.2, whereas traditional low skill-intensive sectors, such as Apparel, Leather, Footwear and Food, are characterized by constant returns.[6] Finally, note that many skill-intensive activities (such as R&D and Marketing) have the nature of fixed costs and therefore tend to generate scale economies.

Moving to our second crucial assumption, several observations suggest that the elasticity of substitution among goods with different skill-intensity is greater than one. A unit elasticity would imply constant expenditure shares over time, but this is contradicted by US data. Between 1970 and 1994, the expenditure share (relative to total manufacturing) in the less skill-intensive textile-apparel-footwear sectors

[6] More precisely, simple calculations on their results show that manufacturing sectors with strong evidence of increasing returns have an average index of skill-intensity (the normalized ratio of workers who completed high school to those who did not) equal to 0.4, while those with constant returns have an average value of 0.12. The remaining sectors, with non-robust estimates of returns to scale, lie in the intermediate range, with an average skill-intensity of 0.23.

has fallen by more than 30%, whereas in modern skill-intensive sectors such as office machinery, pharmaceuticals and electrical machinery it has risen by 160%, 100% and 50%, respectively.[7] More interestingly, in our model the elasticity of substitution in production between skilled and unskilled workers coincides with the elasticity of substitution in consumption between goods of different skill intensity.[8] We can then refer to studies that provide estimates of the former parameter. Freeman [1986] concludes his review of the empirical evidence suggesting a value of the elasticity of substitution between more and less educated labor in the range between 1 and 2. Hamermesh and Grant [1979] review 20 estimates of the elasticity of substitution between production and non-production workers and find a mean estimate of 2.3. Finally, using a different macroeconomic approach, Krusell et al. [2000] report an estimate of 1.67 for the US economy, while Katz and Murphy [1992] find a value of 1.41.

We also confront our result with the data, by considering a panel of countries observed in the years 1970, 1980 and 1990. Our model suggests that the skill premium is increasing in openness to trade and country size, and decreasing in the endowment of skilled workers. We therefore regress the skill premium on the ratio of imports plus exports to GDP, the size of the labor force and the share of workers with secondary education. The results are consistent with the predictions of our model. The coefficients of these variables have the expected sign and are highly significant. Ceteris paribus, a doubling of the degree of openness is associated with a 30% increase in the skill premium; a doubling of country size is associated with a 7% increase in the skill premium, and a doubling of the share of workers with secondary education leads to a 20% fall in the skill premium. These results seem robust with respect to the method of estimation and the specification of the regression equation.

We are not alone in reconsidering the role of international trade in explaining wage inequality. Neary [2001] and Thoenig and Verdier [2001] develop models where trade liberalization between similar countries can lead to skill-biased technical

[7]The source of the data used to calculate these figures is the OECD STAN Database.

[8]More in general, in the Appendix we show that the aggregate elasticity of substitution between factors is a function of those in the individual sectors, together with that between sectors in demand; we then show that an aggregate elasticity of substitution in production greater that one also implies an elasticity in consumption greater than one.

change. The idea underling their models is that of "defensive innovation": increased competition makes skill-intensive technologies more profitable because they deter the entry of new firms. In contrast, we show that even abstracting from technical change and strategic considerations, the trade-induced expansion in market size is sufficient to increase inequality. Our result is also related to Acemoglu [1999] and Acemoglu and Zilibotti [2001]. In their view, North-South trade induces skill-biased technical change by making skill-complement innovations more profitable. However, trade between identical countries plays no role and trade opening in a developing country is unlikely to have an effect on the direction of technical change, since no single developing country has the economic size to affect world incentives. Finally, Manasse and Turrini [2001] show that, in the presence of heterogeneity among skilled workers, trade can spur within-group wage inequality. Our contribution to this growing literature is to consider an asymmetry in scale economies that is both empirically relevant and able to reconcile several puzzling facts. We therefore believe that our mechanism, so far neglected, can be important in understanding how trade affects skill premia.[9]

The plan of the paper is as follows. Section 2 illustrates the basic model, analyzes the effects of international trade on the skill premium and shows that the intersectoral mobility of physical capital may magnify the skill-biased scale effect. Section 3 reconciles the role of trade in explaining wage inequality with the main stylized facts. Section 4 tests the implications of the model using data from a panel of countries. Section 5 concludes.

2 A Simple Model

2.1 Preferences

Consider a country endowed with H units of skilled workers and L units of unskilled workers, where two final goods are produced. Consumers have identical homotetic

[9] An alternative approach, taken by Ethier [2002], is to disregard sectoral asymmetries to focus instead on the intra-sectoral substitution between inputs. Ethier shows that trade and technical progress can increase wage inequality provided that skilled labor and equipment are complement and that unskilled labor and outsourcing are substitutes.

preferences, represented by the following CES utility function:

$$U = \left[(Y_l)^{\frac{\varepsilon-1}{\varepsilon}} + (Y_h)^{\frac{\varepsilon-1}{\varepsilon}} \right]^{\frac{\varepsilon}{\varepsilon-1}}, \tag{3.1}$$

where Y_h and Y_l stand for the consumption of final goods h and l, respectively, and $\varepsilon > 1$ is the elasticity of substitution between the two goods. The relative demand for the two goods implied by (3.1) is:

$$\left(\frac{P_h}{P_l} \right)^{-\varepsilon} = \frac{Y_h}{Y_l}, \tag{3.2}$$

where P_h and P_l are the final prices of goods l and h, respectively. Note that $\varepsilon > 1$ implies that a fall in the relative price induces a more than proportional increase in relative demand. This is a crucial assumption for our results.

2.2 Production and Market Structure

Goods h and l are produced by perfectly competitive firms by assembling n_i $(i = l, h)$ own-industry differentiated intermediate goods. In particular, we assume that the production functions for final goods take the following CES form:[10]

$$Y_i = \left[\int_0^{n_i} y_i\left(v\right)^{\frac{\sigma_i-1}{\sigma_i}} dv \right]^{\frac{\sigma_i}{\sigma_i-1}}, \tag{3.3}$$

where $y_i\left(v\right)$ is the amount of the intermediate good type v used in the production of good i, and σ_i is the elasticity of substitution among any two varieties of intermediates used in sector i. In the following, we assume that $\sigma_l > \sigma_h > \varepsilon$. In words, the elasticity of substitution among intermediates is greater in sector l than in sector h. Further, the elasticity of substitution in production among intermediates used in each sector is greater than the elasticity of substitution in consumption between the final goods.

[10]As discussed later on, these production functions exhibit increasing returns to scale and were introduced into trade theory by Ethier (1982).

The price for final good i (equal to the average cost) implied by (3.3) is:

$$P_i = \left[\int_0^{n_i} p_i \left(v \right)^{1-\sigma_i} dv \right]^{1/(1-\sigma_i)}, \tag{3.4}$$

where $p_i \left(v \right)$ is the price of the intermediate good type v used in the production of good i.

The two sectors producing intermediates are monopolistically competitive *a là* Dixit and Stiglitz [1977] with symmetric firms. The production of each intermediate in sector i involves a fixed requirement, F_i, and a constant marginal requirement, c_i, of labor. In order to keep the algebra as simple as possible, we assume that the two sectors are extreme in terms of skill-intensity, so that sector h uses only skilled workers H, whereas sector l uses only unskilled workers L. In the Appendix, we generalize our results to a setting where both sectors use both types of labor. Hence, the total cost function of a single variety produced in sector i is:

$$TC_i = (F_i + c_i y_i) w_i, \tag{3.5}$$

where w_h and w_l are the wage rates of skilled and unskilled workers, respectively.

Profit maximization by intermediate goods firms in the two sectors implies a markup pricing rule:

$$p_i(v) = p_i = \left(1 - \frac{1}{\sigma_i} \right)^{-1} c_i w_i = w_i, \tag{3.6}$$

where the latter equality follows from a choice of units such that $c_i = \left(1 - \frac{1}{\sigma_i}\right)$. Hence, we have:

$$\frac{p_h}{p_l} = \omega, \tag{3.7}$$

where $\omega = w_h/w_l$ is the skill premium. Intuitively, the relative price of any variety of sector h intermediates is an increasing function of the skill premium, since h is skill-intensive relative to l.

A free-entry condition guarantees zero profits in equilibrium:

$$\pi_i(v) = \pi_i = \left(\frac{y_i}{\sigma_i} - F_i \right) w_i = 0$$

and hence

$$y_i = F_i \sigma_i = 1, \tag{3.8}$$

where the latter equality follows from setting $F_i = 1/\sigma_i$.[11]

Equations (3.6) and (3.8) allow us to simplify the expressions for P_i and Y_i:

$$Y_i = n_i^{\frac{\sigma_i}{\sigma_i - 1}} \tag{3.9}$$

$$P_i = n_i^{\frac{1}{1-\sigma_i}} p_i. \tag{3.10}$$

As equation (3.9) shows, the elasticity of Y_i with respect to n_i is greater the lower is σ_i. Hence, σ_i can be interpreted as an inverse measure of sectoral scale economies. Our assumption $\sigma_l > \sigma_h$ is thus equivalent to assuming stronger increasing returns to scale in sector h than in sector l.[12]

2.3 General Equilibrium

Conditions for full employment of skilled and unskilled workers determine the number of varieties produced in each sector:

$$n_l = L \quad \text{and} \quad n_h = H. \tag{3.11}$$

[11] This assumption is meant to simplify the algebra only and is innocuous for the purpose of the paper. As shown in the next footnote, our normalizations do not affect the *elasticity* of the skill premium to a change of any parameters (they only affect its *level*).

[12] A production function $Y = f(v)$ exhibits scale economies if $f(\lambda v) > \lambda f(v)$ for $\lambda > 1$. An index of scale economies is the elasticity of $f(\lambda v)$ with respect to λ: $\frac{\partial f(\lambda v)}{\partial \lambda} \frac{\lambda}{f(\lambda v)} = \frac{\sigma_i}{\sigma_i - 1}$. This index is clearly decreasing in σ_i. Note also that this measure does *not* depend on marginal and fixed costs: without our simplifying assumptions on F_i and c_i, the general expression for sectoral output with full employment is

$$Y_i = \sigma_i F_i^{1/(1-\sigma_i)} \left(\frac{i}{1 + c_i \sigma_i} \right)^{\sigma_i/(\sigma_i - 1)}, \qquad i = L, H$$

and the elasticity of output to a change in the number of employed workers i is still $\frac{\sigma_i}{\sigma_i - 1}$. This also proves that our normalizations on F_i and c_i affect the level of output but not its scale elasticity. Since our main results depend on the scale elasticity of output, they are unaffected by the normalizations.

Let $\theta = H/\overline{L}$ be the country share of skilled workers in the total workforce, $\overline{L} = H + L$. Equations (3.11) can then be rewritten as:

$$n_l = (1-\theta)\,\overline{L} \quad \text{and} \quad n_h = \theta\overline{L}. \tag{3.12}$$

Substituting (3.9), (3.10), (3.7) and (3.12) into (3.2), and rearranging gives an equilibrium expression for the skill premium:

$$\left[\theta\overline{L}\right]^{\frac{\sigma_h-\epsilon}{\epsilon(\sigma_h-1)}}\omega = \left[(1-\theta)\,\overline{L}\right]^{\frac{\sigma_l-\epsilon}{\epsilon(\sigma_l-1)}}. \tag{3.13}$$

2.4 Trade and the Skill Premium

We can now analyze the effects of trade integration on the skill premium. Since we focus on equilibria with factor price equalization (FPE), we can obtain the free trade prices by applying the above results to a hypothetical integrated economy whose endowments are the sum of those of each country. In particular, totally differentiating equation (3.13) and using the implicit function theorem, we can decompose the change in the skill premium into the following components:

$$\frac{d\omega}{\omega} = \left[\frac{(\epsilon-1)(\sigma_l-\sigma_h)}{\epsilon\,(\sigma_h-1)\,(\sigma_l-1)}\right]\frac{d\overline{L}}{\overline{L}} - \left[\frac{\sigma_h-\epsilon}{\epsilon\,(\sigma_h-1)} + \frac{\sigma_l-\epsilon}{\epsilon\,(\sigma_l-1)}\frac{\theta}{1-\theta}\right]\frac{d\theta}{\theta}. \tag{3.14}$$

Equation (3.14) shows how the skill premium is affected by a variation in the size of the economy ($d\overline{L}/\overline{L}$) and the relative scarcity of skilled workers ($d\theta/\theta$). We use equation (3.14) to first study the effect of intra-industry trade on wage inequality. As shown by Krugman [1979], in a Dixit-Stiglitz framework trade integration among two identical countries is formally equivalent to an increase in country size, $\overline{L}$. Given that $\sigma_l > \sigma_h > \epsilon > 1$, equation (3.14) implies that the coefficient of $\frac{d\overline{L}}{L}$ is positive, and that its magnitude depends positively on the elasticity of substitution ϵ and the sectoral asymmetries $(\sigma_l - \sigma_h)$ in the degree of returns to scale. Thus, pure intra-industry trade among identical countries, often presumed to have no distributional effects, turns out to be skill-biased.

Equation (3.14) also shows the effect of inter-industry trade on wage inequality. Integration between dissimilar countries still implies an increase in the overall size of the economy, but will also change the perceived relative scarcity of factors. Since the coefficient of $d\theta/\theta$ is negative, an increase (fall) in the relative supply of skilled labor has the effect of reducing (increasing) the skill premium.[13] This effect works through the well-known mechanics of the Heckscher-Ohlin-Samuelson theorem, and can dampen or magnify the upward pressure on the skill premium due to the market size effect.[14]

What drives the skill bias of trade? From (3.10), an equi-proportional increase in n_l and n_h lowers the relative price of the skill-intensive good, since this sector enjoys stronger economies of scale. By expanding market size, international trade thus raises the relative productivity of the skill-intensive good, an effect equivalent to skill-biased technical change. Moreover, since the elasticity of substitution in consumption is greater than one, the share of the skill-intensive good in total output rises after trade integration. It follows that trade increases the share earned by skilled labor in total income and hence the skill premium.

2.5 Introducing Physical Capital

We now show how the introduction of physical capital, assumed to be mobile across sectors, magnifies the skill-biased scale effect of trade. With physical capital (K), the total cost function of a single variety produced in sector i becomes:

$$TC_i = (F_i + c_i y_i) r^{\gamma} w_i^{1-\gamma}, \tag{3.15}$$

where r is the rental rate and γ is the share of capital in sector i total cost. For simplicity, and without loss of generality, equation (3.15) considers the case where capital intensity is the same in both sectors ($\gamma = \gamma_h = \gamma_l$). The relative price of skill-intensive varieties implied by (3.15) and profit maximization becomes:

$$\frac{p_h}{p_l} = \frac{r^{\gamma} w_h^{1-\gamma}}{r^{\gamma} w_l^{1-\gamma}} = \omega^{1-\gamma}. \tag{3.16}$$

[13]Note that the coefficient of $d\theta/\theta$ is negatively affected by the elasticity of substitution ϵ, as a high substitutability implies a weak price effect of an increase in the relative supply.

[14]Equation (3.14) can be interpreted as the elasticity of factor prices with respect to the factor content and volume of trade relative to endowments. See Deardorff (2000) and Deardorff and Staiger (1988) for a similar formulation without the scale effect.

Equations (3.2), (3.9) and (3.10) are unchanged; together with (3.16) they imply:

$$n_h^{\frac{\sigma_h-\epsilon}{\epsilon(\sigma_h-1)}}\omega^{1-\gamma}=n_l^{\frac{\sigma_l-\epsilon}{\epsilon(\sigma_l-1)}}. \tag{3.17}$$

Using Shephard's lemma, the demand for each factor can be found from the total cost function (3.15). Noting that $\frac{\partial}{\partial w_i}r^\gamma w_i^{1-\gamma}=(1-\gamma)r^\gamma w_i^{-\gamma}$ and $\frac{\partial}{\partial r}r^\gamma w_i^{1-\gamma}=\gamma r^{\gamma-1}w_i^{1-\gamma}$, we have that the conditions for full employment of physical capital, skilled and unskilled workers are given by:

$$\begin{aligned} K &= \gamma r^{\gamma-1}w_h^{1-\gamma}n_h(F_h+c_hy_h)+\gamma r^{\gamma-1}w_l^{1-\gamma}n_l(F_l+c_ly_l) \\ H &= (1-\gamma)r^\gamma w_h^{-\gamma}n_h(F_h+c_hy_h) \\ L &= (1-\gamma)r^\gamma w_l^{-\gamma}n_l(F_l+c_ly_l). \end{aligned} \tag{3.18}$$

After setting $w_l=1$, we can use (3.18) to express n_h and n_l as functions of the skill premium and the exogenous variables:

$$n_h=\frac{H\omega^\gamma}{(1-\gamma)^{1-\gamma}}\left(\gamma\frac{L+H\omega}{K}\right)^{-\gamma} \quad \text{and} \quad n_l=\frac{L}{(1-\gamma)^{1-\gamma}}\left(\gamma\frac{L+H\omega}{K}\right)^{-\gamma}. \tag{3.19}$$

Substituting (3.19) into (3.17) and solving for ω gives the equilibrium skill premium. Differentiating with respect to ω, K and $\overline{L}=H+L$, and using the implicit function theorem, we find the elasticities of the skill premium to changes in the scale of the economy to be:

$$\frac{d\omega}{\omega}=\frac{\left[\gamma\frac{dK}{K}+(1-\gamma)\frac{d\overline{L}}{\overline{L}}\right]\frac{(\epsilon-1)(\sigma_l-\sigma_h)}{\epsilon(\sigma_h-1)(\sigma_l-1)}}{1-\gamma\left[\frac{\epsilon-1}{\epsilon}\frac{1}{1-\theta+\theta\omega}\left(\frac{\sigma_h(1-\theta)}{\sigma_h-1}+\frac{\sigma_l\theta\omega}{\sigma_l-1}\right)\right]}, \tag{3.20}$$

where again $\theta=H/\overline{L}$ is the share of skilled workers in the total labor force.[15] Note that the coefficient multiplying the scale variables in the square bracket of the numerator is equal to the scale elasticity in (3.14). But now the denominator in (3.20) is less than one and decreasing in γ.[16] Therefore, the effect on the skill premium

[15]The elasticity to a change in the relative skill-endowment θ is here omitted, though straightforward to calculate, because we are interested in showing how capital reallocation affects the scale effect.

[16]Note that, assuming decreasing marginal returns to capital in both sectors, we have $\gamma\frac{\sigma_i}{\sigma_i-1}<1$ for $i=h,l$. This ensures that the denominator of (3.20) is positive.

of trade integration among two identical countries, i.e., a doubling of both K and $\overline{L}$, is now greater, the greater is the share γ of capital in the total cost. Further, equation (3.20) shows that capital accumulation and capital inflows tend to increase the skill premium, as they contribute to expand the scale of the economy. This result is consistent with the literature documenting capital-skill complementarities (see Krusell et al. [2000], among others). To see why capital magnifies the effects of trade integration on the skill premium, it is instructive to study the change in the allocation of capital between the two sectors:[17]

$$\frac{K_h/n_h}{K_l/n_l} = \omega^{1-\gamma}. \tag{3.21}$$

Equation (3.21) shows that the trade-induced rise in the skill premium is associated with a relative increase in capital intensity of firms operating in sector h. The reason is that, by expanding market size, trade integration increases the relative productivity of the resources used in the sector enjoying stronger economies of scale. Hence, trade implies an increase in the relative marginal productivity of capital in sector h. Since in equilibrium the rental rate must be equalized between the two sectors, the only way of restoring the equality after trade integration is by shifting capital out of the less skill-intensive sector and into the skill-intensive sector. As a consequence, the endowment of capital per worker rises for the skilled and falls for the unskilled, which further increases wage inequality.

A similar mechanism is at work in Caselli [1999], where a skill-biased technological revolution induces a reallocation of capital toward the skill-intensive sectors. He also provides evidence of a substantial increase in the US sectoral dispersion of capital intensities since the mid-seventies. In particular, Caselli documents that capital flew to skill-intensive industries during the period of rising wage inequality. Our contribution is to show that such a reallocation of capital can be also due to trade integration. Therefore, capital mobility magnifies the effects of trade integration on the skill premium and strengthens the quantitative relevance of our analysis.

[17] To find (3.21), note that $K_i r = \gamma P_i Y_i$ then use (3.9), (3.10) and (3.16).

3 Trade and Wages: Reconsidering the Facts

In this section, we show how our model can reconcile an important role of trade in explaining the rising skill premia with the main stylized facts. The first critique to traditional trade-based explanations concerns their quantitative relevance: North-South trade flows simply do not seem to be large enough to significantly affect wage premia. Compared to the standard Heckscher-Ohlin approach, our model is less exposed to this criticism as it shows that the entire volume of world trade matters for inequality and not only its net factor content. It remains to argue that the trade-induced skill-biased scale effect might be of significant magnitude. To do so, we compute the scale elasticity of the skill premium given by equation (3.20). A conventional value for the capital share, γ, is 1/3. In the model we use, the elasticity of substitution between goods of different skill-intensity, ϵ, is the same as the elasticity of substitution between skilled and unskilled workers. As already mentioned in the Introduction, estimates of the latter elasticity are in the range (1.5 - 2). We therefore choose the benchmark values of 1.5 and 2. For the remaining parameters, we refer to Antweiler and Trefler [2002] who find that skill-intensive sectors have an average scale elasticity around 1.2 and that traditional less skill-intensive sectors show no departure from constant returns. We therefore set $\sigma_l = \infty$ and let σ_h vary. To have a sense of the economic meaning of σ_h, note that it can be interpreted as an inverse index of the degree of increasing returns. In particular, the production function (3.9) implies that a scale elasticity of output in the range from 1.1 to 1.2 corresponds to a value of σ_h between 11 and 6.

The result is depicted in Figure 3.2, which shows the scale elasticity of the skill premium (on the vertical axis) as a function of σ_h (on the horizontal axis).[18] The broken line corresponds to $\epsilon = 2$, whereas the solid one represents the case with $\epsilon = 1.5$. Figure 3.2 can be used to perform simple experiments. For example, with $\sigma_h = 6$ (a value often used in trade models to describe manufacturing sectors and consistent with several studies[19]) and $\epsilon = 1.5$, the graph shows that the elasticity of the skill premium to the scale of the economy implied by the model is around

[18]Note from equation (3.20) that $d\omega/\omega$ also depends on θ and ω. Numerical simulations show their effect to be negligible. To draw Figure 2, we have used values of 0.35 and 1.4, respectively.

[19]See, for example, Feenstra [1994] and Lai and Trefler [1999].

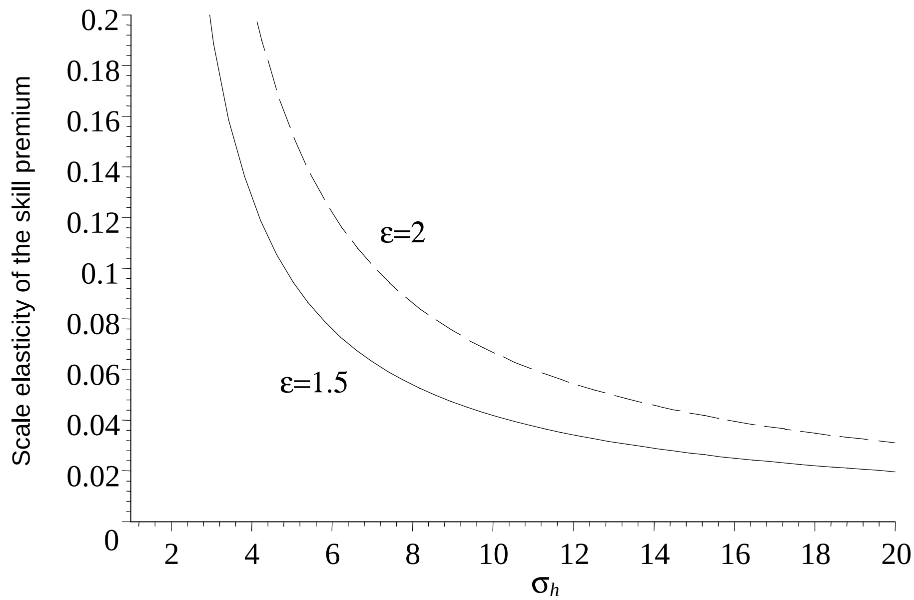

Figure 3.2: Scale elasticity of the skill premium

0.08. In this case, trade integration between two identical countries would imply a 8% increase in the skill premium. More generally, Figure 3.2 shows that even small asymmetries in the sectoral returns to scale are enough to produce a significant effect of market size on wage inequality. This simple quantification suggests that empirical studies focusing on North-South trade might be missing an important mechanism through which globalization enhances skill premia.

A second observation, seemingly at odds with trade models, is that commercial liberalizations in developing countries seem to be followed by increases in wage premia (e.g., Hanson and Harrison [1999] and Robbins [1996]). Our model can rationalize this fact if the skill-biased scale effect is strong enough to overcome the factor proportions effect in skill-scarce countries. To see whether this is more than just a theoretical possibility, we use our model to study the episode of trade integration between Mexico and the United States. This case is of particular interest, because prior to 1985 Mexico could be considered a closed economy due to heavy policies of trade protection. In 1985, Mexico announced its decision to join the GATT and undertook major reforms leading to a reduction in tariffs by 45% and import licenses by more than 75% within three years. During the same period, the skill premium, starting from a value of 1.84, rose by at least 17%. The Mexico experience is also interesting because its major trade partner is the skill-abundant United States. We then perform the following thought experiment. Assume that Mexico was in autarky in 1985; what does our model say about the effect of a complete and instantaneous trade integration with the United States? Using data[20] for the manufacturing sector and the share of white-collar workers as a measure of skilled labor, we have that a move from autarky to the integrated equilibrium implies for Mexico a 4.8 fold increase in the total labor force, a 10.1 fold increase in the aggregate capital stock and a 28.4% increase in the share of white-collar workers. Using these numbers together with the above mentioned parameter values ($\gamma = 1/3$, $\epsilon = 1.5$, $\sigma_h = 6$, $\sigma_l = \infty$), our model predicts the following change in the Mexican skill premium:

$$\frac{d\omega}{\omega} = +0.49 - 0.27 = 0.22$$

where the first number represents the positive scale effect and the second number the negative factor proportions effect. Overall, trade opening in skill-scarce Mexico can lead to an impressive 22% increase of the skill premium. We recognize that this number is notably too high, due to the extreme nature of our exercise. Still, its magnitude suggests that the market size effect can play a significant role in developing countries that experience drastic trade liberalizations.

The third puzzling fact that a satisfactory model should explain is the evolution of relative prices. Though the empirical findings are sometimes mixed, they tend to suggest a *decline* in the relative price of skill-intensive goods during the period of rising skill premia. Our model can help understand this evidence, as it breaks the simple positive relation between good prices and factor prices of standard trade theory. On the one hand, a trade-induced expansion in market size lowers the

[20] Berman, Bound and Griliches [1994] provide the share of US white-collar workers. The equivalent share for Mexico is reported in Hanson and Harrison [1999]. The labor force in manufacturing is taken from the World Development Indicators. The total capital stock is computed from the Penn World Tables.

relative final price of the skill-intensive good:

$$\frac{P_h}{P_l} = \left[\frac{n_l^{\frac{\sigma_l}{\sigma_l - 1}}}{n_h^{\frac{\sigma_h}{\sigma_h - 1}}} \right]^{1/\varepsilon}.$$

Our assumption $\sigma_l > \sigma_h$ implies that a larger market is associated with a lower relative price of the skill-intensive final good: as the skill-intensive sector is characterized by stronger returns to scale, its output grows more after an increase in the market size and this depresses its relative price. On the other hand, trade increases the relative price of each variety of intermediates in the skill-intensive sector, together with the skill premium, because of the stronger productivity gain:

$$\frac{p_h}{p_l} = \omega^{1-\gamma}.$$

These contrasting implications concerning the effects of international trade on price indexes and prices of individual goods may shed light on the mixed results emerging from empirical studies using different methodologies and different levels of sectoral aggregation. In particular, it is suggestive that Lawrence and Slaughter [1993] show a decline in the relative price of skill-intensive goods using a high level of aggregation, whereas Krueger [1997] finds an increase using highly disaggregated data.

4 Empirical Evidence

4.1 The Determinants of Skill-Premia in a Panel of Countries

We now want to confront some empirical implications of our model with data from a panel of countries observed in the years 1970, 1980 and 1990. As shown in section 2, in the case of complete integration and FPE equilibria, the skill premium does not depend on country characteristics, but only on the size and the endowment of the world economy. More generally, in the presence of some kind of trade barriers, our model suggests that the skill premium is increasing in a country's size and openness to trade, and decreasing in the share of skilled workers in its labor force. We have therefore collected data on skill-premia, trade openness, measures of country size and skill endowments. Unfortunately, international data on wages for different categories of workers are difficult to find. Following other empirical studies, we have used the U.N. General Industrial Statistics database to compute the skill premium as the ratio of nonproduction to production (operatives) wage in total manufacturing.[21] Due to the limitations of this dataset, our sample comprises 35 countries in the years 1980 and 1990, and only 26 in 1970[22]. The countries in our sample are at various stages of economic development (from Ethiopia to the United States), with an average real GDP per capita equal to 41% of the US value in 1990. We have then regressed the log(skill premium) on log(openness), measured as the ratio of imports plus exports to GDP, log(labor force) and log(secondary schooling), measured as the share of workers with at least secondary education.[23] The main results are reported in Table 1.

In the first column, estimation is by pooled OLS with a dummy for Latin America. The three variables of interest have the expected sign and are highly significant. Ceteris paribus, a doubling of the degree of openness is associated with a 30% increase in the skill premium; a doubling of the scale is associated with a 7% increase in the skill premium,[24] and a doubling in the share of workers with secondary education leads to a 20% fall in the skill premium. Finally, note that the dummy for Latin America is positive, highly significant and of very large magnitude: being a

[21]To our knowledge, these are the best availabe data for international comparisons on skill premia (see Berman et al. [1998] on this).

[22]The list of countries is the following (a star (*) indicates no data available for the year 1970): Australia, Austria, Bangladesh, Canada, Chile, Colombia, Cyprus*, Czechoslovakia, Denmark, Egypt*, Ethiopia*, Finland, Germany (West), Greece, Guatemala, Hungary, India, Ireland, Italy, Japan, Korea (Republic), Malaysia*, Malta*, Mexico*, Spain*, Sweden, Tanzania, Pakistan*, Panama, Peru, Philippines, Turkey, United Kingdom, United States of America, Uruguay*. Note: the skill premium for Mexico goes back to 1986 only. Due to data availability, the other Mexican observations for 1980 are replaced by those in year 1985. Futher, as far as the skill premium in 1970 is concerned, in the case of Turkey, Peru and Guatemala we use data relative to 1969, 1972 and 1973, respectively.

[23]Data on openness are from the Penn World Tables (Mark 5.6). Labor Force is taken from the World Development Indicators (World Bank). The educational attainment of the total population aged 25 and above is provided by Barro-Lee (School attainment for Ethiopia is available for 1995 only. In the case of Tanzania, school attainment is proxied by the average for sub-Saharan countries).

[24]This estimated value for the scale elasticity of the skill premium is in line with the quantitative implications of the model discussed in Section 3.

Latin American country implies, ceteris paribus, a higher skill premium by more than 40%. This result is in line with other studies on the determinants of inequality, where a dummy for Latin America generally features prominently.[25]

Table1. Skill premia across countries

Variable	Pooled OLS	Random-effects
log(Openness)	0.304	0.271
	(0.103)	(0.061)
log(Labor force)	0.073	0.059
	(0.027)	(0.025)
log(Secondary schooling)	-0.202	-0.192
	(0.049)	(0.039)
Dummy: Latin America	0.425	0.408
	(0.069)	(0.085)
Number of observations	96	96
R-squared	0.53	0.52

Notes: the dependent variable is log(Skill premium). Standard errors are reported in parentheses. In the Pooled OLS regression, robust standard errors are calculated in the presence of repeated observations on individual countries.

Figures 3.3 and 3.4 provide a graphical representation of the partial relation between the skill premium and our two scale variables, log(openness) and log(labor force). The vertical axis reports the value of the skill premium after partialling out the estimated effects of the variables other than openness (Figure 3), and labor force (Figure 4). Looking at the plots, it is apparent that the regression lines are not driven by outliers and that they fit well even widely different observations (such as those for India and Malta).[26]

[25] See, for instance, Barro [2000]. We have also tried with continent dummies for Europe, Africa, Asia and North America, but they turn out not significant.

[26] In order to test more rigorously for the presence of influential observations, we have computed the dfbetas for the coefficients of the scale variables (the difference between the regression coefficients when each observation is included and excluded, scaled by the standard error of the coefficients). We have found that only one observation (Malaysia in 1990) shifts the estimates by more than 0.5 standard errors. Omitting this observation from the sample leaves the significance of our results unchanged.

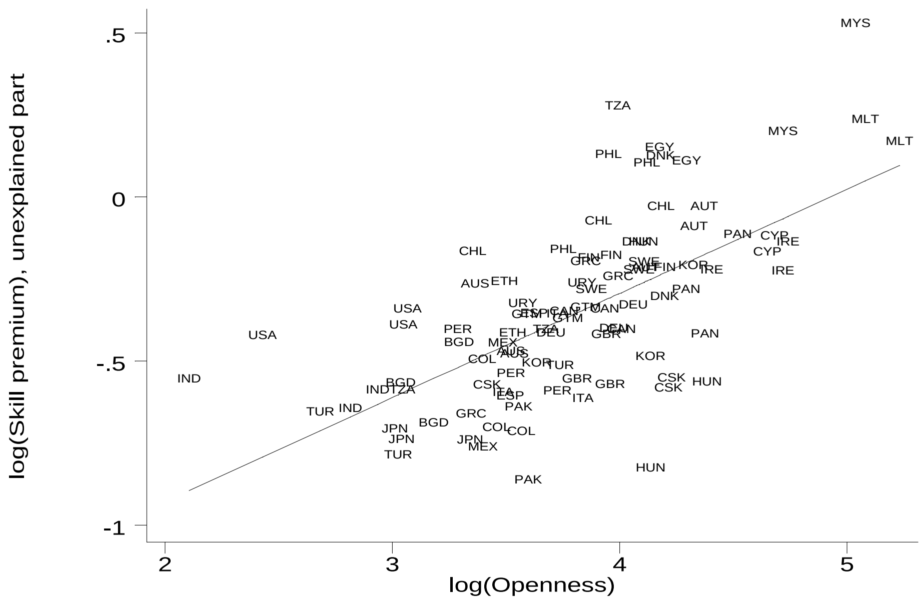

Figure 3.3: Skill premia and trade openness

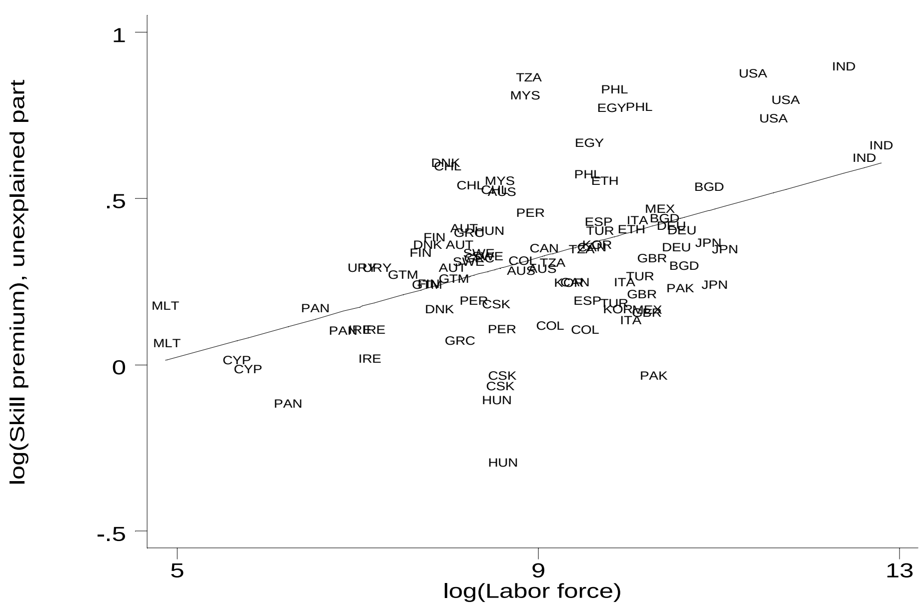

Figure 3.4: Skill premia and country size

To better exploit the limited time variation in the sample, the second column reports random-effects GLS estimates. Once more, all the regressors are highly significant and of the same order of magnitude, although the coefficients are slightly lower than the previous estimates. The Hausman specification test is not significant, with a *p-value* of 0.21, suggesting that individual effects are uncorrelated with the regressors, and hence the appropriateness of the random-effects estimator.[27] Our estimates can be used to assess the impact on wage inequality of the evolution of our scale variables over the period 1970-1990. For instance, the estimates reported in the second column of Table 1 imply that the growth of the U.S. labor force (+37%) and of U.S. trade openness (+64.8%) can account for an almost 20% increase in the skill premium over the two decades of analysis. This simple exercise suggests that scale effects can be important in explaining not only the cross-section of skill-premia, but also the time-series experience of each country.

Table 2 checks the robustness of our results with respect to the proxies for our variables of interest and the specification of the regression equation. In the first and fifth columns we use college attainment instead of secondary schooling as a proxy for education. The results are similar, although the size and significance of the coefficients of interests are slightly reduced. In the second and sixth columns we use total GDP instead of labor force as a proxy for domestic scale. This alternative scale variable has the expected sign but is not highly significant. Although total GDP captures well the economic size of a country, its correlation with per capita GDP makes it a less attractive scale proxy. As shown in the third and seventh columns, per capita GDP has a negative and significant impact on the skill premium; controlling for it restores the significance of total GDP. Finally, in the fourth and eight columns we check the robustness of the basic regression equation with respect to a number of potentially relevant variables. We add time dummies and control for log(per capita GDP) and its square, i.e., for a Kuznets-type relation between wage inequality and per capita GDP. The coefficients on log(openness), log(labor force)

[27]When using the within-estimator, only openness preserves its significance (with a coefficient of 0.222 and a standard error of 0.080). Note, however, that the time variation of labor force and education is poorly measured. This implies that the within-estimator, which only uses temporal variation to estimate coefficients, may yield very imprecise estimates for the impact of these variables on wage inequality.

and log(schooling) still have the expected sign and their size and significance are not affected.

Table 2. Sensitivity analysis

Variable	Pooled OLS				Random-effects			
Openness	.241	.240	.281	.340	.224	.242	.266	.322
	(.098)	(.103)	(.104)	(.116)	(.063)	(.061)	(.062)	(.069)
Labor force	.065		-	.073	.053	-	-	.058
	(.030)			(.032)	(.027)			(.028)
GDP	-	.043	.065	-	-	.040	.057	-
		(.028)	(.028)			(.026)	(.026)	
Sec. schooling	-	-.244	-.134	-.146	-	-.217	-.149	-.142
		(.068)	(.087)	(.085)		(.049)	(.060)	(.067)
College Education	-.114	-	-	-	-.101	-	-	-
	(.033)				(.029)			
GDP p.c.	-	-	-.122	-.816	-	-	-.099	-.872
			(.059)	(.400)			(0.051)	(.470)
GDP p.c. squared	-	-	-	.047	-	-	-	.050
				(.026)				(.028)
Latin America	.467	.381	.402	.483	.446	.382	.392	.470
	(.065)	(.071)	(.067)	(.080)	(.095)	.087	(.084)	(.096)
1990	-	-	-	-.023	-	-	-	-.031
				(.060)				(.056)
Dummy: 1980	-		-	-.095	-	-	-	-.103
				(.054)				(.047)
Observations	96	96	96	96	96	96	96	96
R-squared	.45	.50	.54	.58	.45	.49	.54	.58

Notes: the dependent variable is log(Skill premium). Standard errors are in parentheses. In the Pooled OLS regression, robust standard errors are calculated in the presence of repeated observations on individual countries.

Independent of the method of estimation, the dummy for 1990 is insignificant whereas that for 1980 is significant with a negative sign. Note, also, that the Kuznets curve is significant at the 10-percent level, although with the unexpected sign. One possible explanation for the negative sign of the coefficient of log(per capita GDP) is that this regressor is highly correlated with the proxies for education (the coefficient of linear correlation between log(per capita GDP) and log(secondary schooling) equals 0.85). This is confirmed by the fact that the inclusion of this term generally reduces the size and significance of the schooling proxy. In contrast, the positive sign of the square of log(per capita GDP) suggests that the negative correlation between skill premia and the share of educated workers weakens among rich countries (in our sample this relationship never reverts to a positive one).[28]

4.2 Evidence from other studies

Other empirical studies lend indirect support to our results. Antweiler and Trefler [2002], using trade data for 71 countries and 5 years, show that a rise in output tends to increase the relative demand for skilled workers. Historical evidence seems consistent with a skill-biased scale effect too: Lindert and Williamson [2001], for example, show that inequality widened during globalization booms and after massive immigration, whereas it decreased in the period 1914-1950 of protectionism and in the presence of massive emigration. Finally, Hine and Wright [1998] report indirect evidence in support of the mechanism illustrated in the paper. With reference to the United Kingdom, they estimate the magnitude of trade-induced productivity effects. Their most interesting result is that trade with other OECD countries has a much stronger effect on productivity than trade with developing countries. This is consistent with our model, *in primis*, because the economic size of the OECD countries (and therefore the trade generated scale effect) is much larger than that of developing countries; *in secundis*, because the UK trade with advanced countries is mainly intra-industry trade in skill-intensive goods characterized by strong scale economies (therefore the more pronounced productivity gain).

[28] When using the within-estimator to estimate regressions in Table 2, openness is always positive and highly significant, whereas education and labor force are generally insignificant.

5 Concluding Remarks

The most original result of our analysis is to show that the scale of an economy can be a key determinant of wage inequality. This is a general result which applies to different contexts. In this paper, we have focussed on the role played by a trade-induced scale effect, instead of country-specific scale effects, such as factor accumulation or technical progress. A first reason for this focus is policy relevance. Trade is the only scale variable that can change abruptly as a consequence of policy reform. Second, if globalization goes far enough, factor prices will mainly be determined at the world level and country-specific variables will lose their importance. Third, trade is fundamental in our story because the scale effect operates through the increase in the number of available intermediates made possible by some form of trade. Finally, our framework shows that a "new trade theory" explanation based on intra-industry trade may reconcile the increase in wage inequality with the empirical evidence often used to discredit more traditional trade explanations. We consider this as an important result per se. Our empirical findings lend support to this choice of emphasis, as they suggest that trade is a major determinant of wage inequality in a panel of countries.

We conclude by discussing our results relative to the main alternative explanations for rising skill premia: models of directed technical change and models of outsourcing. In models of directed technical change (initiated by Acemoglu [1998]), a market size effect plays a key role: the skill bias of the technology depends positively on the size of the relative endowment of skilled workers.[29] Since there are no sectoral asymmetries, however the skill premium can only be affected by asymmetric changes in the economic environment. Our mechanism is complementary, as it shows that once sectoral asymmetries are introduced, even a purely symmetric endowment shock affects income distribution.

In models of outsourcing it is instead the relocalization in developing countries of production by OECD countries (through trade in intermediates) that increases

[29] Other evidence in support of the role of technology over time is within-industry demand shifts in favor of skilled labor. Katz and Murphy [1992], among others, document this phenomenon for the United States. See Aghion [2002] for a discussion of various implications of growth models on the dynamics of inequality.

the demand for skilled labor.[30] This happens because the outsourced activities are unskilled-labor intensive relative to those performed in the developed world, but skilled-labor intensive relative to those performed in the developing countries. Our approach shares the basic insight that trade in intermediate inputs does not only affect the import-competing sector, but also the input-using sector. Despite this similarity, the two models apply to quite different situations, as outsourcing typically takes place between dissimilar countries, whereas we emphasize the role of trade in intermediates among industrial countries.

We have derived our results for a specific market structure (monopolistic competition) and specific functional forms on the basis of our reading of the empirical evidence, to have a sense of the quantitative significance of the effect we discuss. Much of debate on trade and inequality is, in fact, centered on the magnitude of the trade-induced effects. But our model is a specific example of a more general principle, surprisingly neglected in the debate: with sectoral asymmetries in the returns to scale and a non-unitary elasticity of substitution in consumption, any increase in market size due to trade integration is non-neutral to income distribution as long as the skill intensity differs across sectors.

6 References

Acemoglu, Daron (1998). "Why Do New Technologies Complement Skills? Directed Technical Change and Wage Inequality," *Quarterly Journal of Economics* 113, 1055-1090.

Acemoglu, Daron (1999). "Pattern of Skill-Premia," NBER Working Paper W 7018.

Acemoglu, Daron and Fabrizio Zilibotti (2001). "Productivity Differences," *Quarterly Journal of Economics* 116, 563-607.

Aghion, Philippe (2002). "Schumpeterian Growth Theory and the Dynamics of Income Inequality," *Econometrica* 70, 855-882.

[30] For a survey, see Feenstra and Hanson [2001].

Antweiler, Werner and Trefler Daniel (2002). "Increasing Returns and All That: A View from Trade," *American Economic Review* 92, 93-119.

Autor, David H., Lawrence F. Katz and Alan B. Krueger (1998). "Computing Inequality: Have Computers Changed the Labor Market?," *Quarterly Journal of Economics* 113, 1169-1213.

Barro, Robert J. (2000). "Inequality and Growth in a Panel of Countries," *Journal of Economic Growth* 5, 5-32.

Berman, Eli, John Bound and Stephen Machin (1998). "Implications of Skill-Biased Technological Change: International Evidence," *Quarterly Journal of Economics* 113, 1245-1280.

Berman, Eli, John Bound and Zvi Griliches (1994). "Changes in the Demand for Skilled Labor within U.S. Manufacturing Industries: Evidence from the Annual Survey of Manufacturing ," *Quarterly Journal of Economics* 109, 367-398.

Caselli, Francesco (1999). "Technological Revolutions," *American Economic Review 89, 78-102.*

Deardorff, Alan V. (2000). "Factor prices and the factor content of trade revisited: what's the use?" *Journal of International Economics* 50, 73-90.

Deardorff, Alan V. and Robert W. Staiger (1988) "An interpretation of the factor content of trade" *Journal of International Economics* 24, 93-107.

Dixit, Avinash and Joseph Stiglitz (1977). "Monopolistic Competition and Optimum Product Diversity," *American Economic Review* 67, 297-308.

Doms, Mark, Timothy Dunne and Kenneth Troske (1997). "Workers, Wages and Technology," *Quarterly Journal of Economics* 112, 253-290.

Ethier, Wilfred (2002). "Globalization, Globalisation: Trade, Technology and Wages," *Tinbergen Institute Discussion Paper* No. 02-088/2.

Ethier, Wilfred (1982). "National and International Returns to Scale in the Modern Theory of International Trade," *American Economic Review* 72, 389-405.

Feenstra, Robert C. (1994). "New Product Varieties and the Measurement of International Prices," *American Economic Review* 84, 157-177.

Feenstra, Robert C. and Gordon H. Hanson (1996). "Foreign Investment, Outsourcing and Relative Wages," in R.C. Feenstra, G.M. Grossman and D.A. Irwin eds., *The Political Economy of Trade Policy: Papers in Honor of Jagdish Bhagwati*, MIT Press, 89-127.

Feenstra, Robert C. and Gordon H. Hanson (1999). "The Impact of Outsourcing and High-Technology Capital on Wages: Estimates for the United States, 1979-1990," *Quarterly Journal of Economics* 114, 907-940.

Feenstra, Robert C. and Gordon H. Hanson (2001). "Global Production Sharing and Rising Inequality: A Survey of Trade and Wages," NBER Working Paper W 8372.

Freeman, Richard B. (1986). "Demand for Education," Chapter 6 in Orley Ashenfelter and Richard Layard (editors) *Handbook of Labor Economics*, North Holland, Vol I, 357-386.

Hall, Robert E. (1986). "Market Structure and Macroeconomic Fluctuations," *Brookings Papers on Econmic Activity* 2, 285-338.

Hamermesh, D. and J. Grant (1979). "Econometric Studies of Labor-Labor Substitution and Their Implications for Policy," *Journal of Human Resources* 14, 518-542.

Hanson, Gordon H. and Ann Harrison (1999). "Trade Technology and Wage Inequality: Evidence from Mexico" *Industrial and Labor Relations Review* 52, 271-288.

Hine, Robert and Peter Wright (1998). "Trade with Low Wage Economies, Employment and Productivity in UK Manufacturing," *The Economic Journal* 108, 1500-1510.

Katz, Lawrence and Kevin Murphy (1992). "Changes in Relative Wages: Supply and Demand Factors," *Quarterly Journal of Economics* 107, 35-78.

Krueger, Alan B. (1997). "Labor Market Shifts and the Price Puzzle Revisited," NBER Working Paper W 5924.

Krugman, Paul (1979). "Increasing Returns, Monopolistic Competition and International Trade," *Journal of International Economics* 9, 469-479.

Krugman, Paul (2000). "Technology, Trade and Factor Prices," *Journal of International Economics* 50, 51-71.

Krusell, Per, Lee E. Ohanian, José-Víctor Ríos-Rull and Giovanni L. Violante (2000). "Capital-skill Complementarity and Inequality: A

Macroeconomic Analysis" *Econometrica* 68, 1029 -1053.

Lai, Huiwen and Daniel Trefler (1999). "The Gains From Trade: Standard Errors with the CES Monopolistic Competition Model," mimeo, University of Toronto.

Lawrence, Robert and Matthew Slaughter (1993). "International Trade and American Wages in the 1980s: Giant Sucking Sound or Small Hiccup?" *Brookings Papers of Economic Activity: Microeconomics* 2, 161-211.

Lindert, Peter H. and Jeffrey G. Williamson (2001). "Does Globalization Make the World More Unequal?" NBER Working Paper w8228.

Manasse, Paolo and Alessandro Turrini (2001). "Trade, Wages and Superstars," *Journal of International Economics* 54, 97 - 117.

Neary, Peter (2001). "Foreign Competition and Wage Inequality," mimeo, University College Dublin.

Paul, Catherine J. Morrison and Donald S. Siegel (1999). "Scale Economies and Industry Agglomeration Externalities: A Dynamic Cost Function Approach", *American Economic Review* 89, 272-290.

Robbins, Donald (1996). "HOS Hits Facts: Facts Win; Evidence on Trade and Wages in the Developing World," HIID Discussion Paper *557*, Harvard University.

Thoenig, Mathias and Thierry Verdier (2001). "A Theory of Defensive Skill Biased Innovation and International Trade," mimeo.

Wood, Adrian (1994). *North-South Trade, Employment and Inequality: Changing Fortunes in a Skill Driven World*, Clarendon Press, Oxford, UK.

Wood, Adrian (1998). "Globalisation and the Rise in Labour Market Inequalities," *The Economic Journal* 108, 1463-1482.

7 Appendix

7.1 The General Model

We study now the more general case in which each good is a Cobb-Douglas composite of H, L and K. We assume that the total cost function of a single variety produced in sector i is:

$$TC_i = (F_i + c_i y_i) r^{\gamma} (w_h^{\alpha_i} w_l^{1-\alpha_i})^{1-\gamma}, \tag{3.22}$$

where r is the rental rate, γ is the share of capital in total cost, and α_i $(i = h, l)$ is the wage-bill share of skilled workers in sector i. We assume that $\alpha_h > \alpha_l$, namely that sector h is skill-intensive relative to sector l. The relative price of skill-intensive varieties implied by (3.22) and profit maximization becomes:[31]

$$\frac{p_h}{p_l} = \frac{r^{\gamma}(w_h^{\alpha_h} w_l^{1-\alpha_h})^{1-\gamma}}{r^{\gamma}(w_h^{\alpha_l} w_l^{1-\alpha_l})^{1-\gamma}} = \omega^{(1-\gamma)(\alpha_h - \alpha_l)}. \tag{3.23}$$

Free-entry and the simplifying assumption $F_i = 1/\sigma_i$ fix the scale of production of each variety to one: $y_i = 1$. Equations (3.2), (3.9) and (3.10) are unchanged;

[31] Prices are a markup over marginal cost, and we have again used the normalization $c_i = \left(1 - \frac{1}{\sigma_i}\right)$.

together with (3.23) they imply:

$$n_h^{\frac{\sigma_h-\epsilon}{\epsilon(\sigma_h-1)}}\omega^{(1-\gamma)(\alpha_h-\alpha_l)}=n_l^{\frac{\sigma_l-\epsilon}{\epsilon(\sigma_l-1)}}. \tag{3.24}$$

The demand for each factor can be found using Shephard's lemma on the total cost function (3.22). After setting w_l as the numeraire, the conditions for full employment of capital, skilled and unskilled workers become:

$$\begin{aligned}
K &= \gamma r^{\gamma-1}\omega^{(1-\gamma)\alpha_h}n_h+\gamma r^{\gamma-1}\omega^{(1-\gamma)\alpha_l}n_l\\
H &= (1-\gamma)\alpha_h r^{\gamma}\omega^{(1-\gamma)\alpha_h-1}n_h+(1-\gamma)\alpha_l r^{\gamma}\omega^{(1-\gamma)\alpha_l-1}n_l\\
L &= (1-\gamma)\left(1-\alpha_h\right)r^{\gamma}\omega^{(1-\gamma)\alpha_h}n_h+(1-\gamma)\left(1-\alpha_l\right)r^{\gamma}\omega^{(1-\gamma)\alpha_l}n_l.
\end{aligned}$$

Solving for n_h and n_l gives:

$$\begin{aligned}
n_i &= \frac{\left(1-\alpha_j\right)H\omega-\alpha_j L}{(1-\gamma)(\alpha_i-\alpha_j)\omega^{\alpha_i(1-\gamma)}}\left(\frac{\gamma}{1-\gamma}\frac{L+H\omega}{K}\right)^{-\gamma}\\
&= \frac{\overline{L}^{1-\gamma}K^{\gamma}\left[\left(1-\alpha_j\right)\theta\omega-\alpha_j(1-\theta)\right]\left(1-\theta+\theta\omega\right)^{-\gamma}}{(1-\gamma)^{1-\gamma}\gamma^{\gamma}(\alpha_i-\alpha_j)\omega^{\alpha_i(1-\gamma)}},
\end{aligned}$$

for $i,j=l,h$, $i\neq j$, $\overline{L}=H+L$ and $\theta=H/\overline{L}$. Simple derivation yields:

$$\frac{\partial n_h}{\partial\omega}>0,\frac{\partial n_l}{\partial\omega}<0,\frac{\partial n_h}{\partial\theta}>0,\frac{\partial n_l}{\partial\theta}<0. \tag{3.25}$$

These partial derivatives come from the production side of the economy. They imply that the higher the supply of one factor, the larger the size of the sector which uses that factor intensively, and that the larger the size of one sector, the higher the relative reward for the factor which is used intensively in that sector. Using the expressions for n_h and n_l in (3.24) and differentiating it with respect to θ, K and $\overline{L}$, we find the elasticity of the skill premium:

$$\frac{d\omega}{\omega}=\frac{\frac{(\epsilon-1)(\sigma_l-\sigma_h)}{(\sigma_h-1)(\sigma_l-1)}\left[\gamma\frac{dK}{K}+(1-\gamma)\frac{d\overline{L}}{\overline{L}}\right]-\left(\frac{\sigma_h-\epsilon}{\sigma_h-1}\frac{\partial n_h}{\partial\theta}\frac{\theta}{n_h}-\frac{\sigma_l-\epsilon}{\sigma_l-1}\frac{\partial n_l}{\partial\theta}\frac{\theta}{n_l}\right)\frac{d\theta}{\theta}}{(1-\gamma)\left(\alpha_h-\alpha_l\right)\epsilon+\frac{\sigma_h-\epsilon}{\sigma_h-1}\frac{\partial n_h}{\partial\omega}\frac{\omega}{n_h}-\frac{\sigma_l-\epsilon}{\sigma_l-1}\frac{\partial n_l}{\partial\omega}\frac{\omega}{n_l}}.$$

Given the inequalities in (3.25) and our assumption $1 < \epsilon < \sigma_h < \sigma_l$, it can be seen that the skill premium is increasing in the scale and decreasing in the share of skilled workers. Equations (3.14) and (3.20) are all special cases of this formula.

Finally, it is possible to show that the elasticity of substitution between skilled and unskilled workers (holding the other variables constant) is given by:

$$\varepsilon_w = -\frac{d(H/L)}{d\omega}\frac{\omega}{H/L}\bigg|_{n_h,n_l,K_h,K_l} = \frac{(\alpha_h - \alpha_l)(\epsilon - 1)}{\left(\frac{\alpha_h}{1-\alpha_h}\frac{L}{H\omega} - 1\right)^{-1} + \left(1 - \frac{\alpha_l}{1-\alpha_l}\frac{L}{H\omega}\right)^{-1}} + 1.$$

Rearranging, we can write the elasticity of substitution in consumption (ϵ) as a function of the elasticity of substitution in production (ε_w):

$$\epsilon = 1 + \frac{(\varepsilon_w - 1)}{\alpha_h - \alpha_l}\left[\left(\frac{\alpha_h}{1-\alpha_h}\frac{L}{H\omega} - 1\right)^{-1} + \left(1 - \frac{\alpha_l}{1-\alpha_l}\frac{L}{H\omega}\right)^{-1}\right].$$

Note that $\varepsilon_w > 1$ implies $\epsilon > 1$ and that $\varepsilon_w = \epsilon$ if $\alpha_h = 1$ and $\alpha_l = 0$, as in the model with extreme factor intensities in the main text.

Chapter 4

Scale and Inequality*

1 Introduction

The theoretical literature has identified three main culprits for the dramatic world-wide increase in wage inequality: skill-biased technical change, capital-skill complementarity and international trade. In this paper, we suggest the existence of an important link among these explanations, so far neglected in the debate. We show that, under plausible assumptions, scale is skill-biased. Therefore, since technical change, as well as factor accumulation and trade integration all imply an expansion in market size, they are essentially skill-biased phenomena, even in the absence of technology biases, complementarities among the inputs or Stolper-Samuelson effects.

To make the point, we formulate a simple two-sector general equilibrium model where firms producing sector-specific intermediates engage in Cournot competition. Our choice of industrial structure is justified by the attractiveness of its properties more than the realism of its assumptions: Cournot is the simplest oligopolistic model where the degree of competition is endogenous and varies in a plausible way with market size. Similarly to Krugman (1979), it implies that a scale expansion involves a pro-competitive effect which forces firms to lower mark-ups and expand output. This is its key property for our purpose.

To keep a high level of generality, we allow the two sectors to differ in skill-intensity, fixed and marginal costs, and the degree of substitutability between inter-

* Written with Paolo Epifani. We thank Daron Acemoglu, Jaume Ventura, Fabrizio Zilibotti and seminar partecipants at IIES, Stockholm University, for comments. Any remaining errors are our own.

mediates. Sector-specific intermediates are aggregated into consumption goods at no cost. Finally, on the demand side, the elasticity of substitution in consumption between the two final goods is taken to be greater than one. Under these assumptions, any increase in market size (due to, for instance, factor augmenting technical progress, factor accumulation or trade integration) brings about an increase in the skill premium as long as skill-intensive sector is less competitive than the other. Interestingly, this will be the case, even in the absence of any asymmetries in technology or demand, if the skilled workers represent a minority of the labor force (true, almost by definition). More in general, we show that the skill-bias of scale is stronger the lower the share of skilled workers in the total workforce, the greater the relative importance of plant-level fixed costs and the lower the relative degree of substitutability among skill-intensive intermediates. The reason for this result is that the pro-competitive effect, coupled with plant-level fixed costs, generates increasing returns to scale that are higher the lower the level of competition in a sector. This is a natural implication of oligopolistic models approaching perfect competition as market size tends to infinity. As a consequence, a less competitive sector becomes relatively "more productive" with an increase in market size and its income share expands because of gross-substitutability in consumption.

A few recent papers address similar issues from different perspectives. Neary (2001) shows that in the presence of oligopolistic markets, increased competition encourages strategic over-investment by incumbent firms in order to deter entry. This raises the ratio of fixed to variable costs and, given that fixed costs are assumed to be skill-intensive, also the skill premium. While interesting, the mechanism illustrated in the paper is quite specific, since it only applies to a few sectors dominated by large firms that interact strategically. In Ekholm and Midelfart-Knarvik (2001), firms can choose between two rather ad hoc technologies: one exhibits high skill-intensive fixed costs and low unskill-intensive marginal costs, while the other exhibits low fixed costs and high marginal costs. They then show that a trade-induced expansion in market size raises the relative profitability of the skill-intensive technology, thereby raising the skill premium. Beyond the low level of generality, a limit of these models where fixed costs are skill intensive is that they tend to have counterfactual implications on mark-ups. Dinopulous et al. (2001) argue that if, as in Krugman (1979), increased competition expands firm size, and if firm size is skill-biased, then trade raises the skill premium. No micro-foundation for the skill bias of firm size is however provided in the paper. In Acemoglu (1999) and Acemoglu and Zilibotti (2001), North-South trade induces skill-biased technical change by making skill-complement innovations more profitable. In these models, however, North-North trade and more generally market size, have no impact on wage inequality. Finally, in Epifani and Gancia (2002) we show that in the presence of an elasticity of substitution in consumption greater than one and stronger increasing returns in the skill-intensive sector, trade integration, even among identical countries, is skill-biased. We also provide evidence in support of our main assumptions. However, the model in the paper does not provide a micro-foundation for the sectoral asymmetries in the scale elasticity of output.

In summary, we add to the literature on the determinants of wage inequality by illustrating a mechanism which, although very simple, is surprisingly more general than the existing ones, since it applies not only to trade-induced increases in market size but to any scale expansion. Further, and most important, it does not rely on ad hoc assumptions on technology, but rather provides a micro-foundation for why skill-intensive sectors become more productive as an economy grows. In the next Section, we formulate a simple model which clarifies our argument. Section 3 discusses the empirical relevance of the main assumptions and implications of the model. Section 4 concludes.

2 The Model

Consider a country endowed with H units of skilled workers and L units of unskilled workers, where two final goods are produced. Consumers have identical homotetic preferences, represented by the following CES utility function:

$$U = \left[\gamma \left(Y_h\right)^{\frac{\epsilon-1}{\epsilon}} + (1-\gamma)\left(Y_l\right)^{\frac{\epsilon-1}{\epsilon}}\right]^{\frac{\epsilon}{\epsilon-1}}, \tag{4.1}$$

where Y_h and Y_l stand for the consumption of final goods h and l, respectively, and $\epsilon > 1$ is the elasticity of substitution between the two goods. γ is a parameter that captures the relative importance of the skill-intensive good in consumption. The relative demand for the two goods implied by (4.1) is:

$$\frac{P_h}{P_l} = \frac{\gamma}{1-\gamma}\left[\frac{Y_l}{Y_h}\right]^{1/\epsilon}, \tag{4.2}$$

where P_h and P_l are the final prices of goods l and h, respectively. Note that $\epsilon > 1$ implies that a fall in the relative price induces a more than proportional increase in relative demand. This is a crucial assumption for our results.

Goods h and l are produced by perfectly competitive firms by assembling at no cost own-industry differentiated intermediate goods. In particular, we assume that in each sector there is a continuum of intermediates of measure one and that the production functions for final goods take the following CES form:

$$Y_i = \left[\int_0^1 Y_i(v)^{\frac{\sigma_i-1}{\sigma_i}}\,dv\right]^{\frac{\sigma_i}{\sigma_i-1}}, \quad i = l, h \tag{4.3}$$

where $Y_i(v)$ is the total amount of the intermediate good type v used in the production of good i, and σ_i is the elasticity of substitution among any two varieties of intermediates used in sector i. The price for final good i (equal to the average cost) implied by (4.3) is:

$$P_i = \left[\int_0^1 p_i(v)^{1-\sigma_i}\,dv\right]^{1/(1-\sigma_i)}, \quad i = l, h \tag{4.4}$$

where $p_i(v)$ is the price of the intermediate good type v used in the production of good i.

Each intermediate v is a homogeneous good produced by a finite number $n_i(v)$ of firms. Firms are symmetric and engage in Cournot competition. The production of each intermediate in sector i involves a fixed requirement, F_i, and a constant marginal requirement, c_i, of labor. Since our main argument holds as long as factor-intensities are different, to keep the algebra as simple as possible we assume that the two sectors are extreme in terms of skill-intensity, so that sector h uses only skilled workers H, whereas sector l uses only unskilled workers L. The total cost function for a producer of intermediate v in sector i is:

$$TC_i(v) = [F_i + c_i y_i(v)]\,w_i, \quad i = l, h \tag{4.5}$$

where y_i is the amount produced by a single firm (because of symmetry, the total amount of intermediate v available in the economy is $n_i(v)y_i(v)$) and w_h and w_l are the wage rate of skilled and unskilled workers, respectively.

Profit maximization by intermediate firms, taking the output of other competing firms as given, implies a mark-up pricing rule:

$$p_i(v) = p_i = \left(\frac{\sigma_i n_i(v)}{\sigma_i n_i(v) - 1}\right) c_i w_i, \quad i = l, h \tag{4.6}$$

A free-entry condition in each industry producing any variety v implies zero profits in equilibrium:

$$\pi_i(v) = \left(\frac{c_i y_i(v)}{\sigma_i n_i(v) - 1} - F_i\right) w_i = 0, \quad i = l, h$$

Full employment requires:

$$[F_h + c_h y_h(v)]\, n_h(v) = H \quad \text{and} \quad [F_l + c_l y_l(v)]\, n_l(v) = L$$

Using this condition together with free entry yields the equilibrium number of firms in each industry and the output produced by each of them:

$$n_h(v) = n_h = \sqrt{\frac{H}{F_h \sigma_h}} \quad \text{and} \quad n_l(v) = n_l = \sqrt{\frac{L}{F_l \sigma_l}} \tag{4.7}$$

$$y_h(v) = y_h = \frac{1}{c_h}\left[\sqrt{H F_h \sigma_h} - F_h\right] \quad \text{and} \quad y_l(v) = y_l = \frac{1}{c_l}\left[\sqrt{L F_l \sigma_l} - F_l\right] \tag{4.8}$$

Note that a scale increase (i.e., an increase in H and L) is associated with a rise in firms' output. This is a direct consequence of the pro-competitive effect of a market size expansion, which reduces price-marginal cost mark-ups and hence forces firms to increase output to cover fixed costs. Note also that, as shown by (4.7), the number of firms grows less than market size. This is the so-called defragmentation effect of a market size expansion (Helpman, 1984): when, due to fiercer competition, the price falls, some firms must exit for the surviving ones to expand their output.

Finally, note that symmetry implies:

$$Y_i = Y_i(v) = n_i y_i, \quad P_i = p_i(v) = p_i. \tag{4.9}$$

The skill-premium (ω) can be found by substituting (4.6), (4.7), (4.8) and (4.9) into (4.2):

$$\omega = \frac{w_h}{w_l} = \frac{\gamma}{1-\gamma}\left(\frac{c_l}{c_h}\right)^{\frac{\epsilon-1}{\epsilon}}\left[\frac{L}{H}\right]^{1/\epsilon}\left[\frac{1-\sqrt{F_h/H\sigma_h}}{1-\sqrt{F_l/L\sigma_l}}\right]^{\frac{\epsilon-1}{\epsilon}}. \tag{4.10}$$

Equation (4.10) shows that the skill premium is higher the higher the relative importance of the skill-intensive good in final consumption, as captured by γ. Further, for $\epsilon > 1$, the skill premium is lower the higher the relative marginal cost of the skill-intensive good (c_l/c_h). With an elasticity of substitution in consumption greater than one, a higher relative marginal cost, by raising the relative price of the final good, reduces the expenditure share on the skill-intensive good and hence also the skill premium. Finally, the term $(L/H)^{1/\epsilon}$ captures the standard factor proportions effect: ceteris paribus, the skill premium is higher the lower the relative supply of skilled workers.

2.1 Scale elasticity of the skill premium

Let $\theta = H/\overline{L}$ be the share of skilled workers in the total workforce, $\overline{L} = H + L$. Equation (4.10) can then be rewritten as:

$$\omega = \left(\frac{\gamma}{1-\gamma}\right)\left(\frac{1-\theta}{\theta}\right)^{1/\epsilon}\left(\frac{c_l}{c_h}\right)^{\frac{\epsilon-1}{\epsilon}}\left(\frac{1-\sqrt{F_h/\theta\overline{L}\sigma_h}}{1-\sqrt{F_l/(1-\theta)\overline{L}\sigma_l}}\right)^{\frac{\epsilon-1}{\epsilon}} \tag{4.11}$$

Differentiating (4.11) with respect to $\overline{L}$ allows to calculate an expression for the scale elasticity of the skill premium:

$$e^{\omega}_{scale} = \frac{d\omega}{d\overline{L}}\frac{\overline{L}}{\omega} = \frac{(\epsilon-1)}{2\epsilon\sqrt{\theta(1-\theta)\overline{L}}}\frac{\sqrt{\frac{(1-\theta)F_h}{\sigma_h}} - \sqrt{\frac{\theta F_l}{\sigma_l}}}{\left(1-\sqrt{\frac{F_l}{(1-\theta)\overline{L}\sigma_l}}\right)\left(1-\sqrt{\frac{F_h}{\theta\overline{L}\sigma_h}}\right)} \tag{4.12}$$

Note, first, that $e^{\omega}_{scale} = 0$ if $\epsilon = 1$. More interestingly, (4.12) shows that, for $\epsilon > 1$, scale is skill-biased, i.e., $e^{\omega}_{scale} > 0$, as long as:

$$\frac{F_h(1-\theta)}{\sigma_h} > \frac{\theta F_l}{\sigma_l} \tag{4.13}$$

This condition can easily be interpreted. Using equations (4.6) and (4.7) it is possible see that (4.13) holds whenever the mark-up over marginal cost is higher in the skill-intensive sector. In other words, an increase in total market size raises wage inequality if the skill-intensive sector is less competitive than the other. A noteworthy implication of inequality (4.13) is that, even in the absence of sectoral asymmetries in fixed costs ($F_h = F_l$) or in the degree of substitutability among intermediates ($\sigma_h = \sigma_l$), scale is skill-biased as long as skilled workers are less than the unskilled, i.e., for $\theta < 1/2$.

The intuition behind this result is the following. In the presence of plant-level fixed costs and variable mark-ups, the pro-competitive and the defragmentation effects associated with market size expansion imply that sectoral production functions are non-homotetic and that returns to scale fall as the size of the market increases. To see this, substitute (4.7) and (4.8) into (4.9) to derive an expression for the sectoral production functions in terms of the model's parameters; then, it is straightforward to show that the scale elasticity of sectoral outputs (e^{Y}_{scale}) is given by:

$$e^{Y}_{scale} = \frac{1-(1/2)\sqrt{F/V\sigma}}{1-\sqrt{F/V\sigma}} \tag{4.14}$$

where we have omitted sectoral indices and $V = H$ or L. Note that the scale elasticity of output is greater than one and decreasing in V. This implies that, for identical fixed costs and degree of substitutability among varieties across sectors, scale efficiency gains from a market size expansion are greater in the skill-intensive sector as long as $H < L$. It follows that output in the smaller and less competitive skill-intensive sector grows more than in the labor-intensive sector after a scale increase, thereby inducing a rise in the skill premium because of gross-substitutability in consumption ($\epsilon > 1$).[1]

[1]The relative output increase in the skill-intensive sector implies that its relative price falls.

Once the internal logic of the model is clear, it is straightforward to see why, as shown by (4.13), the skill-biased scale effect is stronger the greater the relative importance of plant-level fixed costs (F_h/F_l) and the lower the relative elasticity of substitution among varieties (σ_h/σ_l) in the skill-intensive sector. As for the former, for a given H/L, a higher F_h/F_l implies higher mark-ups in the skill-intesive sector and hence larger scale efficiency gains. Similarly, a lower σ_h/σ_l implies that, ceteris paribus, firms in the skill-intensive sector are less competitive, produce less and are subject to a higher scale efficiency gain.

Two other results are worth noting. First, (4.14) shows that, since the scale elasticity of sectoral production functions converges to one for V approaching infinity, the skill bias of scale vanishes when the scale grows very large. This effect is captured by $\overline{L}$ in the denominator of (4.12). Second, while the relative real marginal cost (c_l/c_h) affects the level of the skill premium (see equation (4.10)), it has no effect on its scale elasticity. The reason is that, as long as variable costs are constant returns to scale, they do not alert the degree of non-homoteticity of the production functions, which is the engine of the skill bias of scale. Hence, what matters for scale to be biased is not the sectoral asymmetry in the ratio of fixed to variable costs (i.e., $F_h/c_h > F_l/c_l$), but rather the asymmetry in the ratio of plant-level real fixed costs to factor endowments (i.e., $F_h/H > F_l/L$).

3 The Evidence

This section briefly discusses the empirical relevance of the main assumptions and implications of our model. On the production side, the model assumes plant-level scale economies that decrease with firm size, since they are generated by fixed costs. The recent plant-level evidence supports this assumption. Tybout and Westbrook (1995) use plant-level manufacturing data for Mexico to show that most industries exhibit increasing returns to scale that typically decrease with larger plant sizes. Similarly, Tybout et al. (1991) and Krishna and Mitra (1998) find evidence of a

---

With an elasticity of substitution in consumption greater than one, the demand for skill-intensive goods increases more than proportionally, raising their share of total expenditure and therefore also the skill premium.

reduction in returns to scale in manufacturing plants after trade liberalization in Chile and India, respectively.[2]

Contrary to the Dixit-Stiglitz (1977) monopolistic competition framework, market structure in our model involves variable mark-ups. In this respect, the evidence is compelling. Country studies reported in Roberts and Tybout (1996), which use plant-level manufacturing data for Chile, Colombia, Mexico, Turkey and Morocco, all find that increased competition due to trade liberalization is associated with falling mark-ups. Similar results using a different methodology are found, among others, by Levinshon (1993) for Turkey, Krishna and Mitra (1998) for India, and Harrison (1994) for Cote d'Ivoire.

Under the above assumptions, the model predicts the skill-bias of scale to be stronger the lower the share of skilled workers in the total workforce (θ), the greater the relative importance of plant-level fixed costs (F_h/F_l) and the lower the relative degree of substitutability among intermediates (σ_h/σ_l) in the skill intensive sector. As for the latter two, note that skill-intensive productions often involve complex activities, such as R&D and marketing, that raise both fixed costs and the degree of product differentiation. Regarding the share of skilled workers in the total workforce, we can refer to the Barro-Lee data base to make a crude cross-country comparison. A common practice in the empirical literature on wage inequality is to define skilled workers as those who have college education. Accordingly, using the percentage of the total population aged 25 and over that completed post-secondary education as a proxy for the share of skilled workers, we find that in 2000 it ranges from a minimum of 0.1% in Gambia, to a maximum of 30.3% in the U.S. (with New Zeland ranking second with a share of 16%).

Further, and most important, the model's prediction of sectoral asymmetries in the scale elasticity of output finds support in two recent empirical studies. Antweiler and Trefler (2002), using international trade data for 71 countries and five years, find that skill-intensive sectors, such as Petroleum Refineries and Coal Products, Pharmaceuticals, Electric and Electronic machinery and Non-Electrical Machinery, have an average scale elasticity around 1.2, whereas traditional low skill-intensive sectors, such as Apparel, Leather, Footwear and Food, are characterized by constant

[2] See also Tybout (2001) on this point.

returns. Using a different methodology, Paul and Siegel (1999) estimate returns to scale in US manufacturing industries at the two-digit industry level for the period 1979-1989. Their estimates of sectoral scale economies are strongly positively correlated with the sectoral skill-intensity.[3]

Next, for these asymmetries to matter for wage inequality, the elasticity of substitution between goods produced with different factor-intensities must be greater than one. In this respect, note that a unit elasticity would imply constant expenditures shares over time, which is contradicted by US data. In Epifani and Gancia (2002), we show that between 1970 and 1994 the expenditure share (relative to total manufacturing) in the less skill-intensive textile-apparel-footwear sectors fell by more than 30%, whereas in modern skill-intensive sectors such as office machinery, pharmaceuticals and electrical machinery it rose by 160%, 100% and 50%, respectively. Further, in our model the elasticity of substitution in production between skilled and unskilled workers is equivalent to the elasticity of substitution in consumption between goods of different skill-intensity. Several studies provide estimates of the former parameter and most of them are above one. Freeman (1986) suggests a value of the elasticity of substitution between more and less educated labor in the range between 1 and 2. Hamermesh and Grant (1979) find a mean estimate of 2.3. Lastly, Krusell et al. (2000) and Katz and Murphy (1992) report estimates for the US economy of 1.67 and 1.41, respectively.

Finally, our model predicts the aggregate relative factor demand to be non-homotetic. This provides an explanation for an empirical result found by Antweiler and Trefler (2002), showing that a 1% scale increase brings about a 0.42% increase in the relative demand for skilled workers. This result is also consistent with Denny and Fuss (1993), who find evidence of skill-biased scale effects in their study on the telecommunication industry.

4 Conclusions

We have shown that, under plausible and fairly general assumptions about market structure, preferences and technology, scale tends to be skill-biased. The mechanics

[3]See also Epifani and Gancia (2002) on this point.

of this result can be summarized as follows. In the presence of plant-level fixed costs and variable mark-ups, a market size expansion involves a pro-competitive effect which causes firms to reduce their price and increase their output, and a defragmentation effect which causes a less than proportional increase in the number of firms in each industry. These effects imply that sectoral production functions are non-homotetic and that the scale elasticity of output, a measure of increasing returns, decreases with the size and the level of competition in a sector. Therefore, as long as skill-intensive sector is less competitive than the other, any increase in market size raises productivity relatively more in the skill intensive sector and, given gross-substitubility in consumption, leads to higher wage inequality. Even in the absence of any asymmetries in technology, this will be the case if the skilled workers represent a minority of the labor force (true, almost by definition). More in general, we have shown the skill-biased of scale to be larger the lower the share of skilled workers in the total workforce, the greater the relative importance of plant-level fixed costs in the skill-intensive sector and the lower the relative degree of substitutability among skill-intensive goods.

We have also shown that our main assumptions and results are corroborated by most plant- and industry-level empirical studies.

Our result that scale is skill-biased provides an important link among major explanations for the worldwide rise in skill premia: skill-biased technical change, capital-skill complementarity and international trade. According to the first, inequality rose because recent innovations in the production process, such as the widespread introduction of computers, have increased the relative productivity of skilled workers.[4] In this respect, an important implication of our model is that, independent of the specific features of technological improvements, factor augmenting technical progress may appear skill-biased simply because it raises the total supply of effective labor in the economy and therefore its scale. Similarly, the capital-skill complementarity argument (see Krusell et al. (2000), among others) emphasizes that, since new capital equipment requires skilled labor to operate and displaces unskilled workers, its accumulation raises the relative demand for skilled labor. More

[4]See, among others, Autor et al.(1998) for empirical evidence and Aghion (2002) for theoretical perspectives.

generally, we have shown that, even in the absence of capital-skill complementarity (indeed, even in the absence of physical capital, though straightforward to incorporate), factor accumulation tends to be skill-biased because it expands the scale of production. Finally, it is often argued that North-South trade liberalization may have increased wage inequality in advanced industrial countries through the well-known Stolper-Samuelson effect. However, the Stolper-Samuelson theorem is silent on the distributional effects of North-North (or South-South) trade, which represents the large majority. Our model suggests, instead, that any kind of trade integration, by increasing the market size for goods, is potentially skill biased.

5 References

Acemoglu, Daron (1999). "Patterns of Skill Premia," *NBER Working Paper* W 7018.

Acemoglu, Daron and Fabrizio Zilibotti (2001). "Productivity Differences," *Quarterly Journal of Economics* 116, 563-607.

Aghion, Philippe (2002). "Schumpeterian Growth Theory and the Dynamics of Income Inequality," *Econometrica* 70, 855-882.

Antweiler, Werner and Trefler Daniel (2002). "Increasing Returns and All That: A View from Trade," *American Economic Review* 92, 93-119.

Autor, David H., Lawrence F. Katz and Alan B. Krueger (1998). "Computing Inequality: Have Computers Changed the Labor Market?," *Quarterly Journal of Economics* 113, 1169-1213.

Dinopoulos Elias, Constantinos Syropoulos and Bin Xu (2001). "Intra-Industry Trade and Wage Income Inequality," mimeo, University of Florida.

Dixit, Avinash and Joseph Stiglitz (1977). "Monopolistic Competition and Optimum Product Diversity," *American Economic Review* 67, 297-308.

Ekholm, Karolina and Karen Helene Midelfart-Knarvik (2001). "Relative Wages and Trade-Induced Changes in Technology," *CEPR Discussion Paper* No.

2677.

Epifani, Paolo and Gino A. Gancia (2002). "The Skill Bias of World Trade," *W.P. CESPRI-Bocconi University* No.129.

Feenstra, Robert C. and Gordon H. Hanson (1996). "Foreign Investment, Outsourcing and Relative Wages," in R.C. Feenstra, G.M. Grossman and D.A. Irwin eds., *The Political Economy of Trade Policy: Papers in Honor of Jagdish Bhagwati*, MIT Press, 89-127.

Freeman, Richard B. (1986). "Demand for Education," Chapter 6 in Orley Ashenfelter and Richard Layard (editors) *Handbook of Labor Economics*, North Holland, Vol I, 357-386.

Hamermesh, D. and J. Grant (1979). "Econometric Studies of Labor-Labor Substitution and Their Implications for Policy," *Journal of Human Resources* 14, 518-542.

Hanson, Gordon H. and Ann Harrison (1999). "Trade Technology and Wage Inequality: Evidence from Mexico" *Industrial and Labor Relations Review* 52, 271-288.

Harrison, Ann (1994). "Productivity, Imperfect Competition and Trade Reform," *Journal of International Economics* 36, 53-73.

Katz, Lawrence and Kevin Murphy (1992). "Changes in Relative Wages: Supply and Demand Factors," *Quarterly Journal of Economics* 107, 35-78.

Helpman, Elhanan (1984). "Increasing Returns, Imperfect Markets and Trade Theory," in R.W. Jones and P. B. Kenen (eds.), *Handbook of International Economics*, Vol.1, North Holland, Amsterdam.

Krishna, Pravin and Devashish Mitra (1998) "Trade Liberalization, Market Discipline and Productivity Growth: New Evidence from India" *Journal of Development Economics* 56, 447-462.

Krugman, Paul (1979). "Increasing Returns, Monopolistic Competition and International Trade," *Journal of International Economics* 9, 469-479.

Krusell, Per, Lee E. Ohanian, José-Víctor Ríos-Rull and Giovanni L. Violante (2000). "Capital-skill Complementarity and Inequality: A Macroeconomic Analysis," *Econometrica* 68, 1029 -1053.

Levinsohn, J.(1993). "Testing the Imports-as-Market-Discipline Hypothesis," *Journal of International Economics* 35, 1-22.

Neary, Peter (2001). "Foreign Competition and Wage Inequality," mimeo, University College Dublin.

Paul, Catherine J. Morrison and Donald S. Siegel (1999). "Scale Economies and Industry Agglomeration Externalities: A Dynamic Cost Function Approach", *American Economic Review* 89, 272-290.

Robbins, Donald (1996). "HOS Hits Facts: Facts Win; Evidence on Trade and Wages in the Developing World," *HIID Discussion Paper 557,* Harvard University.

Roberts, Mark J. and James R. Tybout (1996). "*Industrial Evolution in Developing Countries*," Oxford University Press, Oxford.

Tybout, James R.(2001). "Plant-and Firm-Level Evidence on 'New' Trade Theories," *NBER Working Paper* W 8418.

Tybout, James R. and M. Daniel Westbrook (1995). "Trade Liberalization and the Dimensions of Efficiency Change in Mexican Manufacturing Industries," *Journal of International Economics* 39, 53-78.

Tybout, James R., J. De Melo and V. Corbo (1991). "The Effects of Trade Reforms on Scale and Technical Efficiency: New Evidence from Chile," *Journal of International Economics* 31, 231-250.

Chapter 5

Geography, Migrations and Equilibrium Unemployment*

1 Introduction

In this paper we study the effects of trade integration on the regional coevolution of income, migrations and unemployment in a dynamic core-periphery model with limited labor mobility and frictions in the job matching process. Our main aim is to investigate the determinants of the geographic distribution of unemployment and to explain some recent trends observed within European regions.

We focus on three main observations. First, during the last two decades, there has been a slight tendency toward convergence in real per capita income among European regions. For instance, Quah (1996) documents a reduction over time in the cross-sectional spread of relative real incomes. In particular, he finds that the standard deviation of the regional per capita income distribution has declined by 8% between 1980 and 1989.[1] Second, we observe a marked change in the historical evolution of migration flows. European labor mobility was high, both between and within countries, in the period between 1955 and 1970. Thereafter, Europe has experienced a sharp decline of interregional and international labor mobility (see,

* Written with Paolo Epifani. We have benefited from the comments of Francesco Daveri, Henrik Horn, James Markusen, Gianmarco Ottaviano and seminar participants at MIT, IIES, Stockholm University and the CNR Conference (Cagliari 2001). We are especially grateful to Torsten Persson and Fabrizio Zilibotti for helpful discussion and advice. We thank Annika Andreasson for editorial assistance. The usual caveat applies.

[1]See also Barro and Sala-i-Martin (1991).

among others, Faini et al. [1997], and Bentolila [1997]). Third, over the last two decades European regions have experienced a dramatic increase in unemployment disparities. For instance, Puga and Overman (2000) find that, between 1986 and 1996, the Gini index of regional concentration of unemployment rose by 19%. They also find that an index of polarization of regional unemployment increased by 37% in the same period.[2] Hence, a puzzle seems to arise from these stylized facts, as they show falling migrations rates despite growing unemployment differentials, and regional income convergence despite polarization and divergence of regional rates of unemployment.[3]

This paper argues that the interplay between the centripetal and centrifugal forces emphasized by the new economic geography (Fujita et al. [1999]), and imperfections in the job matching process of the kind stressed by the equilibrium unemployment theory (Pissarides [1990]) can be helpful in explaining these facts. Although the analysis of unemployment is virtually absent from models of the new economic geography[4], geographical factors seem to matter a great deal for regional unemployment rates. As an example, Puga and Overman (2000) find that unemployment rates are more similar across neighboring regions, despite international borders, than across regions with similar skill composition or sectoral specialization. This suggests that the geography of access to markets and to the sources of intermediate inputs can explain a relevant fraction of the regional unemployment levels. Furthermore, a recent surge of theoretical and empirical studies highlights the importance of supply side considerations for understanding the labor market outcomes and emphasizes the key role played by unemployment subsidies in explaining the high unemployment rates experienced by European countries. We believe that this literature, which treats unemployment as an equilibrium phenomenon, can also shed

[2]See also, among others, Obstfeld and Peri (1998).

[3]These trends are particularly evident among Spanish regions. Bentolila (1997) reports that net regional migration rates fell by almost 90% between the early sixties and the early nineties. In the same period, an index of absolute unemployment rate differentials across regions increased by almost a fivefold factor. Finally, an index of dispersion of regional per capita income fell by 35% percent in the same period.

[4]More in general, only recently there has been a growing interest in studying the relationships between trade and unemployment. For an analysis of the effects of international trade on unemployment in presence of imperfections in the job matching process see, among others, Davidson, Martin and Matusz (1999) and Jansen and Turrini (1998).

light on the regional evolution of European labor markets.

To combine these two approaches, we build a symmetric two-region model in which trade costs generate agglomeration economies. We introduce a fixed factor and limited labor mobility to capture centrifugal forces and hence to restrain the incentive to concentrate all the economic activities in a single region. Frictions in the job matching process lead to equilibrium unemployment and a search cost modeled in terms of intermediate goods generates a positive externality of agglomeration on the labor market. Finally, we consider a central government that provides a common unemployment subsidy. We then use this model to study the effect of trade integration on income inequality, unemployment differentials and migration flows.

In this paper, regional integration is the engine of regional evolutions. We interpret it in a broad sense so as to include the process of European institutional integration, as well as the improvement in interregional communication networks due to both technical progress and investment in road and telecommunication infrastructure. Finally, as long as regional economies become increasingly weightless (Quah, 1997), a lower share of resources is to be devoted to shipment of goods. This phenomenon is thus isomorphic to a fall in trade costs.

Our main results can be summarized as follows. Starting from a symmetric equilibrium with high trade costs, trade integration triggers a wave of migrations which lead to the emergence of a core-periphery equilibrium, with strong regional disparities both in terms of per capita income and unemployment rates. Thereafter, a further reduction of trade barriers generates income convergence, which in turn reduces the incentive to migrate. At the same time, the unemployment differential still grows larger until a higher level of integration is reached. Therefore, assuming that in the early Eighties the European regions were in a core-periphery equilibrium, the model implies that regional integration leads to convergence in real per capita income, divergence in regional unemployment rates, and declining migration rates. Hence, the model can help explain the most striking features of last decades European regions' unemployment experience.

A second important result of the paper is to highlight a sharp contrast between the short run impact of labor mobility and its effects on the long run equilibrium. In particular, during the transitional dynamics migration tends to promote regional unemployment convergence, whereas in the long run it exacerbates the unemployment differential. We also show how a change in a common policy implemented by a central government can affect in a very different way the two regions, both in the short run dynamics and in the long run.

The last contribution of the paper is methodological. The introduction of search frictions in the job market allows us to study the geographical allocation of production in a fully dynamic framework. This enables us to analyze the stability properties of equilibria in a formal way and to study the dynamic adjustment of the economy after a change in the environment. An interesting by-product of our approach is that it allows to address a common methodological problem in the new economic geography literature. In these models recourse is made to the following ad hoc migration rule: in each period, a fraction of workers moves toward the region that offers a higher real wage and away from the region with a lower real wage.[5] This assumption, which is deprived of deep justification, is necessary in these models in order to avoid catastrophic agglomeration after trade integration. In our model we do not need this assumption, because we can show that, since the matching process between jobs and workers requires time, it implies a gradual relocation of both firms and workers. Hence, we avoid catastrophic agglomeration without imposing any ad hoc assumptions on sluggish labor mobility.

The paper is organized as follows. Section 2 sets out the formal model. Section 3 analyzes the steady-state properties of the model and illustrates the effects of trade integration on regional variables. Section 4 studies the transitional dynamics of the system after a fall in trade costs. Section 5 analyzes a policy experiment. Section 6 concludes.

2 The Model

In this section we describe a core-periphery model along the lines of Krugman (1991). We depart from the literature on economic geography by introducing frictions in the labor market and imperfect labor mobility. We study an economy in which there are two regions, North and South (indexed by $i = N$, S), two sectors, agriculture

[5] See Fujita et al. (1999, page 62).

and manufacturing, and two factors, farmers and workers. The two regions share the same preferences, technology and original endowments. To capture the notion of "distance" between the two regions (to be interpreted in a broad sense), we consider a trade cost on manufactured goods only. The agricultural sector employs farmers to produce an homogeneous good. Firms in manufacturing use workers and intermediates to produce a variety of manufactured goods. We assume that farmers account for a fraction $(1-2\alpha)$ of total population, which is normalized to unity. Farmers are immobile and divided evenly between the two regions.[6] Workers are mobile, but incur in a non-monetary migration cost, which is increasing in the share of immigrants over the labor force. To preserve symmetry, we assume that the number of workers born in each region is equal to α and is constant over time. The final distribution of workers among the two regions is determined endogenously. We introduce equilibrium unemployment among manufacturing workers by assuming frictions in the job matching process. Finally, we consider an unemployment benefit, equal in both regions, financed by a lump sum tax levied on the whole population.

2.1 Households

Consumers have identical Cobb-Douglas preferences over an agricultural good, A, and a composite bundle of differentiated manufactured goods, M. Risk-neutral individuals have time separable preferences, discount future utility at the rate $(1+r)^{-1}$, where r is the interest rate. Time is discrete[7]. In the region of birth, utility is given by:

$$U_i(0) = \sum_{t=0}^{\infty} (1+r)^{-t} \left[M_i(t)\right]^{\mu} \left[A_i(t)\right]^{1-\mu}$$

Utility maximization implies that in each period a fixed share of income, μ, is devoted to manufactured goods.

The composite bundle M is defined as a CES function over a continuum of

[6]The assumption of immobile farmers, which generate local demand for manufactured goods, ensures that, even in the presence of strong forward and backward linkages in manufacturing, there is still an incentive to keep some manufacturing activity in the peripheral region, consisting in the lower competition for local farmers' demand. See Krugman (1991).

[7]In order to save on notation, in the following we omit the time index from all the static equations.

measure n of varieties produced in the whole economy:

$$M_i = \left[\int_0^n m_i(v)^{\frac{\sigma-1}{\sigma}} dv\right]^{\frac{\sigma}{\sigma-1}} \tag{5.1}$$

where $\sigma > 1$ is the elasticity of substitution between any two varieties and $m_i(v)$ represents consumption of variety v in region i. By minimizing the cost of obtaining one unit of M_i we find the price index for the composite bundle:

$$q_i = \left[\int_0^n p(v)^{1-\sigma} dv\right]^{1/(1-\sigma)} \tag{5.2}$$

where $p(v)$ is the final price of variety v. Demand for each variety v is obtained by using Shephard's lemma on the expenditure function $q_i M_i$:

$$m_i(v) = \frac{p(v)^{-\sigma}}{q_i^{1-\sigma}} \mu Y_i$$

where Y_i is total income in region i, and μ is the share of income devoted to the composite bundle implied by Cobb-Douglas preferences, so that $\mu Y_i = q_i M_i$.

Manufacturing workers can migrate, but if they leave the region of birth they incur in a non-monetary cost which reduces their instantaneous utility in every period by a factor $1/\lambda \leq 1$. We assume λ to be increasing in the share of the original work force which leaves the region of birth[8]. For analytical convenience, λ is modeled as a CES function:

$$\lambda_i = \max\left\{\left(\frac{\alpha}{L_i}\right)^{1/\epsilon}, 1\right\}$$

where α and L_i are the original and final labor force in region i, respectively. Note that ε can be interpreted as an index of the degree of labor mobility. Given these

[8]This assumption can be justified on several grounds. Here, we mention two. The first is racism, which may plausibly increase with the share of immigrants in total workforce, thus reducing their welfare. The second relates to the housing market (which we do not explicitly model): as the share of migrants rises, house rents rise in the region of immigration and fall in the region of emigration. Hence, if we assume that emigrants own a house (only) in their region of birth, then their welfare is, ceteris paribus, a decreasing function of the share of emigrants. Alternatively, a realistic assumption would have been individual heterogeneity in migration cost. We have not gone along this this way since it gives rise to uninteresting complications. However, our assumption captures in a reduced-form fashion this heterogeneity.

assumptions, a worker born in region i will migrate to region j if and only if $U_j \geq \lambda_i U_i$. Using the definition of λ, we have that the fraction of the original working population which does not move is:

$$L_i = \min\left\{\alpha\left(\frac{U_i}{U_j}\right)^{\varepsilon}, \alpha\right\} \tag{5.3}$$

for $i, j = N, S$ and $i \neq j$[9]. In the final equilibrium, both employed and unemployed workers will be indifferent between living in the two regions.

2.2 Production and Labor Market

The agricultural sector employs farmers to produce a homogeneous good under constant returns to scale and perfect competition. The agricultural good is freely traded and is taken as the numeraire. Labor productivity is set to one so that the equilibrium wage for farmers is unity. Agriculture is modelled as a residual sector tied to land, its main role being to sustain the demand for goods from peripheral regions. For this reason we interpret it in a broad sense that includes different traditional activities which cannot be easily relocated. For simplicity, we do not study farmers' unemployment[10].

Manufacturing firms produce a large variety of differentiated goods which are used both for final consumption and as intermediate inputs. Firms are symmetric, each of them needs one worker and a fixed amount $1/a$ of the composite bundle M per unit of time. This intermediates requirement captures in a simplified fashion investment in capital equipment and its maintenance. Firms and workers are matched in the labor market through a process that requires time. This assumption captures the idea that heterogeneities in skills and jobs make it costly for a firm or a worker to find a suitable partner. Once employed, a worker produces one unit of a single

[9] Following Faini (1996), we assume that the original working population of each region is constant and is not affected by ongoing migrations; in other words, we consider only a guest-worker type of mobility and we implicitly assume that no offsprings of the guest-worker are born in the visited region. This assumption is not crucial for our results, but allows us to consider the possibility of return migrations, an interesting feature supported by some empirical evidence.

[10] It would be possible to introduce unemployment in agriculture in a way that parallels the manufacturing sector. Under mild assumptions the unemployment rate in the two sectors will evolve in a similar fashion and none of the qualitative results of the model will be affected.

variety which coincides with the final output of the firm. Since the price of any variety is decreasing in the quantity supplied, no two firms will find it convenient to produce the same variety. Furthermore, as differentiated goods can be traded, each region will specialize in a different range of varieties so that $n_N \cup n_S = n$. Given the symmetry in production and demand, every variety from each region will have the same production price p_i. Production prices can differ from final prices because of an "iceberg" trade cost: of $\tau > 1$ units shipped to the other region, only one unit arrives at destination. This implies that the final price in region i of a variety produced in region j is $p_j\tau$ and the price index (5.2) reduces to:

$$q_i = \left[n_i p_i^{1-\sigma} + n_j \left(p_j \tau\right)^{1-\sigma}\right]^{1/(1-\sigma)} \tag{5.4}$$

for $i, j = N, S$ and $i \neq j$.

We now describe the matching process in the regional labor markets, which are assumed to be segmented. As a firm decides to enter the market, it has to post a vacancy and a new job is immediately created. Production will start only once a worker has been found, but in order to keep the job, filled or vacant, a firm must pay the fixed cost q_i/a for intermediates. This assumption reflects the fact that maintaining idle equipment is expensive and makes the search process costly for a firm.[11] Since the price index q_i depends on trade costs, the fixed cost in intermediates creates linkages between firms that make agglomeration of production advantageous. It also provides a parsimonious way to introduce the link between the location of production and the labor market (the search cost now depends on the trade cost), which has been stressed in the early literature on agglomeration economies (see Marshall [1920]). Following Pissarides (1985, 1990), the frictions generated by heterogeneity in the labor market are summarized by a function that gives the measure of successful matches per unit of time. In the simplest approach, this function depends positively on the number of job seekers and the number of vacant jobs. For tractability, we assume that it takes the Cobb-Douglas form $du_i^{\eta} v_i^{1-\eta}$, where u_i represent the unemployment rate, v_i the number of searching firms as a fraction of the labor force and d is a scaling parameter. Defining $\theta_i = v_i/u_i$ as the

[11] Following Pissarides (1985), we assume that equipment can be brought into use, rented and scrapped instantaneously, implying that the number of job vacancies is a perfectly flexible variable.

"tightness" of the labor market, we can write the probability that an unemployed worker will be matched as $d\theta_i^{1-\eta}$. Similarly, the probability that a firm will fill a vacancy is $d\theta_i^{-\eta}$. Matches are destroyed at the exogenous rate s. Upon separation, both the firm and the worker must reenter the labor market.

The value at time t of a firm with a filled job, $V_{fi}(t)$, can be expressed as the sum of its profits at time t, $p_i(t) - w_i(t) - q_i(t)/a$, plus the expected discounted value of the firm at time $t+1$:

$$V_{fi}(t) = p_i(t) - w_i(t) - q_i(t)/a + (1+r)^{-1}[(1-s)V_{fi}(t+1) + sV_{vi}(t+1)] \quad (5.5)$$

where the expression in brackets is the expected value of the firm a time $t+1$. Note that with probability s the match is destroyed, and hence the value of the firm falls to $V_{vi}(t+1)$, which represents the value at time $t+1$ of a searching firm.

Similarly, the value at time t of a firm posting a vacancy, $V_{vi}(t)$, must be equal to the cost of idle equipment, $-q_i(t)/a$, plus the expected discounted value of the firm in the next period:

$$V_{vi}(t) = -q_i(t)/a + (1+r)^{-1}[(1 - d\theta_i^{-\eta}(t))V_{vi}(t+1) + d\theta_i^{-\eta}(t)V_{fi}(t+1)] \quad (5.6)$$

where the expression in brackets is the expected value of a searching firm at time $t+1$. Note that the value of the firm rises to $V_{fi}(t+1)$ in case of a successful match, i.e., with probability $d\theta_i^{-\eta}(t)$.

We assume free entry of firms, hence, the value of posting a vacancy must be zero. Imposing $V_{vi} = 0$ in (5.6) yields:

$$V_{fi}(t+1) = \frac{(1+r)\, q_i(t)/a}{d\theta_i^{-\eta}(t)} \quad (5.7)$$

Using (5.7) into (5.5) and imposing $V_{vi} = 0$, we obtain:

$$V_{fi}(t) = p_i(t) - w_i(t) - (q_i(t)/a)[1 - (1-s)/d\theta_i^{-\eta}(t)] \quad (5.8)$$

The value at time t for an employed worker, $V_{ei}(t)$, equals the wage rate, plus the expected discounted value of the worker at time $t+1$:

$$V_{ei}(t) = w_i\,(t) + (1+r)^{-1}[(1-s)V_{ei}(t+1) + sV_{ui}(t+1)] \tag{5.9}$$

Note that, with probability s the match is destroyed and the value for the worker falls to $V_{ui}(t+1)$, which represents the value for an unemployed worker.

Finally, the value for a job seeker equals:

$$V_{ui}(t) = z(t) + (1+r)^{-1}[(1-d\theta_i(t)^{1-\eta})V_{ui}(t+1) + d\theta_i(t)^{1-\eta}V_{ei}(t+1)] \tag{5.10}$$

where $z(t)$ is an unemployment benefit, equal in both regions, provided by the central government. We assume that z is financed through a lump sum tax, T, levied on the whole population.

Wages are flexible, i.e., there is renegotiation in each period (see Pissarides [1985]). They are determined as the solution to a Nash bargaining problem, implying that the worker surplus is a constant fraction β of the total surplus generated by the match:

$$V_{ei} - V_{ui} = \beta\,(V_{ei} - V_{ui} + V_{fi}) \tag{5.11}$$

Finally, in each period t, $sn(t)$ jobs are exogenously destroyed, whereas $d\theta_i(t)^{1-\eta}u_i(t)L_i(t)$ new jobs are created. Hence, the number of producing firms, which is identically equal to the number of employed workers, evolves according to the following law of motion:

$$n_i(t+1) = (1-s)n_i(t) + d\theta_i(t)^{1-\eta}u_i(t)L_i(t) \tag{5.12}$$

2.3 General Equilibrium

In order to close the model we impose the following general equilibrium constraints. First, regional income is given by farmers' income, equal to $(1-2\alpha)/2$, plus manufacturing wages, equal to n_iw_i, and net manufacturing profits. The last term comprises profits of firms with a filled job, which sum up to $n_i(p_i - w_i - q_i/a)$, minus the losses incurred by $v_iL_i(=\theta_iu_i)$ firms with a job vacancy. Hence we can write:

$$Y_i = (1-2\alpha)/2 + p_in_i - (q_i/a)(n_i + \theta_iu_iL_i) \tag{5.13}$$

Given regional income, market clearing in manufacturing requires the total supply of each variety (one unit) to equal total demand for consumption and intermediate goods from both regions:

$$1 = \frac{p_i^{-\sigma}\mu Y_i}{q_i^{1-\sigma}} + \frac{p_i^{-\sigma}\tau^{1-\sigma}\mu Y_j}{q_j^{1-\sigma}} + \frac{n_i + \theta_i u_i L_i}{a}\left(\frac{p_i}{q_i}\right)^{-\sigma} + \frac{n_j + \theta_j u_j L_j}{a}\left(\frac{p_i}{q_j}\right)^{-\sigma}\tau^{1-\sigma} \quad (5.14)$$

for $i, j = N, S$ and $i \neq j$.

Since we allow for equilibrium unemployment, the labor market clearing condition is replaced by the requirement that the number of employed workers be equal to the number of active firms:

$$n_i = L_i\,(1 - u_i) \quad (5.15)$$

Finally, balanced government budget requires the subsidy to equal the ratio of government revenues from the lump sum tax, T, over the sum of unemployed workers in the two regions:

$$z = \frac{T}{L_i u_i + L_j u_j} \quad (5.16)$$

for $i, j = N, S$ and $i \neq j$.

3 Steady state analysis

In a steady-state all the variables must be constant. Solving equations (5.9) and (5.10) for $V_{ei}(t) = V_{ei}(t+1)$ and $V_{ui}(t) = V_{ui}(t+1)$, we obtain:

$$V_{ei} = \left(\frac{r+1}{r}\right)\frac{sz + (r + d\theta_i{}^{1-\eta})w_i}{(r + s + d\theta_i{}^{1-\eta})} \quad (5.17)$$

$$V_{ui} = \left(\frac{r+1}{r}\right)\frac{(r+s)z + d\theta_i{}^{1-\eta}w_i}{(r + s + d\theta_i{}^{1-\eta})} \quad (5.18)$$

Similarly, imposing $V_{fi}(t) = V_{fi}(t+1)$ in (5.8) and (5.7) gives the following price equation:

$$p_i = w_i + (q_i/a)[1 - (r+s)/d\theta_i^{-\eta}] \quad (5.19)$$

It can be shown that in steady-state the wage equation is given by the following expression:[12]

$$w_i = (1-\beta)\, z + \beta \left[p_i + (\theta_i - 1) \frac{q_i}{a} \right] \tag{5.20}$$

In words, the wage rate compensates a fraction $(1-\beta)$ of the lost unemployment benefit and gives the worker a share β of the firm's output in excess of production costs and of the average vacancy cost per unemployed worker.

Using the wage equation (5.20) into (5.19), we obtain the equilibrium price of a variety produced in region i:

$$p_i = z + \frac{q_i/a}{1-\beta} \left(\beta\theta_i + (1-\beta) + \frac{r+s}{d\theta_i^{-\eta}} \right) \tag{5.21}$$

As a final requirement, in the steady-state the number of unemployed workers is constant. From (5.12), this implies that the flow of laid off workers offsets exactly the flow of job seekers who are hired. Hence, from (5.12) and (5.15), the steady-state rate of unemployment is given by:

$$u_i = \frac{s}{s + d\theta_i^{1-\eta}} \tag{5.22}$$

Summarizing, the steady-state of the system is described by equations (5.3), (5.4), (5.13)-(5.18), (5.20)-(5.22), and by the equivalent equations for region j.

We can now explore the steady-state properties of the model. Since the system is highly non linear, it cannot be solved analytically. Therefore, we proceed by

[12] In order to obtain the wage equation (5.20), set: $V_{fi}(t) = V_{fi}(t+1)$, $V_{ei}(t) = V_{ei}(t+1)$, $V_{ui}(t) = V_{ui}(t+1)$ and $V_{vi} = 0$. Then, use (5.9) and (5.5) into (12). This gives:

$$(r+1)\, w_i = rV_{ui} + \beta\,(r+1)\,(p_i - q_i/a - rV_{ui})$$

In order to eliminate rV_{ui} from the RHS, use (5.7) into (5.11). This gives:

$$V_{ei} - V_{ui} = (1+r)\,\frac{\beta}{1-\beta}\,\frac{q_i/a}{d\theta_i^{-\eta}}$$

Now use the expression for $V_{ei} - V_{ui}$ into (5.10). This gives:

$$rV_{ui} = (1+r)\, z + (1+r)\,\theta(q_i/a)\frac{\beta}{1-\beta}$$

Using the expression for rV_{ui} into the expression for $(1+r)\, w_i$ gives the wage equation (5.20).

numerical simulations. Our main aim is to study the impact of regional integration on regional inequalities, with particular reference to the geographic distribution of unemployment. Hence, in this section we analyze the structure of steady-states as a function of trade costs.

Before turning to numerical examples, we briefly summarize the forces which affect the geographical structure of the economy. Since the two regions are originally identical, the model will always exhibit a symmetric equilibrium in which manufacturing production is evenly distributed. But the presence of labor mobility implies that a geographically differentiated production structure may arise. The specific outcome depends on the migration choice, which is in turn determined by the interaction of two opposing forces, one working toward agglomeration, the other against it. Agglomeration forces consist, *in primis*, of the forward and backward linkages among manufacturing firms and, *in secundis*, of the forward and backward linkages among consumers and producers. These forces attract firms and workers towards the region with the larger market to save on transport costs. Centrifugal forces arise because competition for local farmers' demand is lower in the smaller region and this tends to increase, ceteris paribus, nominal wages and profits in the peripheral region. The existence of increasing migration costs further reduces the incentive for agglomeration. Consistent with a well established result from the new economic geography literature[13], we find that for very high or very low trade costs centrifugal forces prevail, so that the symmetric equilibrium is unique. Conversely, agglomeration forces prevail for intermediate levels of trade costs. In this case, the symmetric equilibrium becomes unstable and a stable core-periphery pattern emerges: workers and firms leave the peripheral region (the South) and manufacturing production becomes partially agglomerated in the core region (the North)[14].

Figure 1 summarizes the steady-state evolution of regional variables as a function of trade costs (from $\tau = 1$ to $\tau = 2$). The parameter values used in these simulations are reported in the appendix. Most of them are taken from other studies (e.g.,

[13]See Ottaviano and Puga (1999) for an analysis of the forces at work in models of the new economic geography.

[14]To analyze local stability properties of equilibria we have linearized the system in a neighborhood of the steady-state. We find that there is always a unique saddle-path stable type of equilibrium (symmetric versus partially agglomerated). Multiple stable equilibria do not arise here because of limited labor mobility. Details on the transitional dynamics are discussed in Section 4.

Pissarides [1998], Fujita et al. [1999]); the remaining ones are chosen to give realistic values for the variables of interest. For instance, these parameters imply that the average duration of a job is about 5.5 years, whereas the average unemployment spell is of 5-6 months. The unemployment benefit varies between 60% and 70% of nominal wages. In all the graphs displayed in Figure 5.1, the solid line represent Northern variables whereas the dashed line refers to the South.

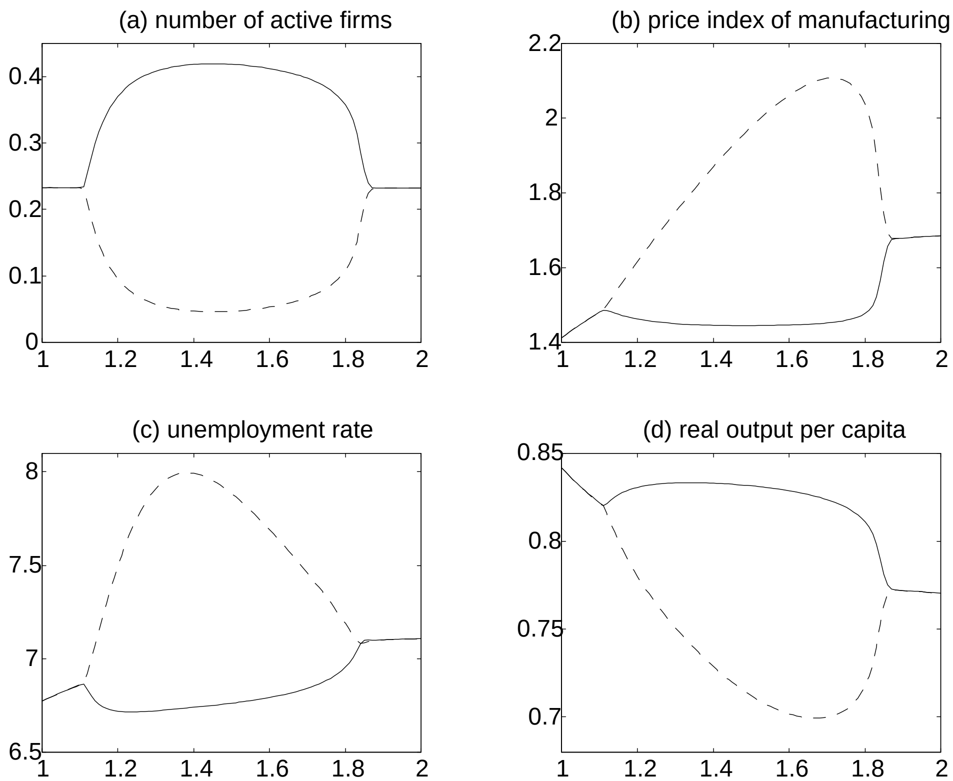

Figure 5.1. steady-states as functions of trade costs

Panel (a) reports the number of firms in the two regions, which equals manufacturing employment. When trade costs are reduced below a threshold level, the symmetric equilibrium breaks down: employment and production agglomerate in the core, although the periphery keeps a small but positive share of manufacturing. The reason for partial agglomeration is that in this model, contrary to Krugman (1991), agglomeration forces are bounded by the increasing costs of migration. Hence, even for intermediate trade costs, i.e., when agglomeration forces are stronger, a positive share of manufacturing workers stays in the periphery. Finally, for low transport costs, the geographical advantage of the core vanishes; the disutility of being an immigrant induces a wave of return migration to the South until the symmetric equilibrium is restored.

Panel (b) reports the price index of manufacturing. When symmetry breaks down, a large mass of workers and firms leave the South, and hence most manufacturing goods must be imported in this region. Consequently, trade costs become a relevant component of the price index. This explains why the South experiences a dramatic increase in the price index of manufacturing. The converse is true in the North, where agglomeration implies a fall in the volume of imports and a consequent fall in the price index of manufacturing. Note, also, that further falls in trade costs imply a different response by the two regions' price indices. Since Northern imports of manufactured goods from the South are very small, the price index is fairly stable in the core region. Conversely, since the South imports most of the manufacturing goods, the fall of its price index closely mirrors the fall of trade costs.

Panel (c) illustrates the evolution of regional unemployment rates (percentages). When the symmetric equilibrium breaks down, the sharp fall of the price index in the North lowers substantially the cost of intermediates and therefore also the search cost, which in turn induces the opening of new vacancies and a rise in the labor market tightness. The opposite happens in the South, where the increase in the price index deteriorates the labor market conditions. This translates into a core-periphery unemployment gap. An interesting feature of this model is that, contrary to conventional wisdom, the unemployment gap is first generated and then exacerbated by migrations. As it will become apparent in the next section, this result holds only in the long run; in fact, during the short run adjustment, migration flows tend to reduce the unemployment gap.

To understand the evolution of the unemployment gap as trade costs are reduced, it is important to study the role played by the subsidy. As symmetry is broken, the higher search cost in the South is partially mitigated by a sharp fall in the real value of the subsidy (the price index for manufacturing increases in the periphery). However, as trade costs fall, the real value of the subsidy grows, and this deteriorates the labor market. Unemployment in the South reaches a peak for intermediate values of transport costs. Thereafter, the negative effect of the growing real subsidy is offset by the fall in the cost of intermediates induced by regional integration, and the Southern unemployment rate starts to decline. Note that, as the geographical advantage of the core vanishes, the periphery experiences a wave of return migrations which reduces the steady-state peripheral unemployment rate (partly at the expense of the North), because it reduces the share of manufacturing goods subject to trade costs. Hence, as migration generated the emergence of regional disparities, return migrations speed up the process of convergence. Finally, note that once the symmetric equilibrium is restored, further falls in trade costs reduce unemployment because they lower the cost of intermediates.

Panel (d) shows the evolution of regional real per capita income. Note that, in the symmetric equilibrium, a fall in trade costs raises per capita income because it reduces unemployment and increases real wages. However, once symmetry is broken, per capita income rises in the North and falls in the South, mainly because of the divergent behavior of the price indexes in the two regions. Further falls in trade costs have little impact on Northern income, since most manufacturing production is concentrated in that region. Conversely, per capita income grows fast in the South, because of the higher real value of wages and subsidies induced by the fall in trade costs.

3.1 Sensitivity analysis

The choice of parameter values does not generally affect the qualitative results illustrated so far. In particular, with regard to "geographic" parameters, a higher share of manufactured goods in consumption, μ, a lower share of farmers in total population (i.e., a higher α), a lower elasticity of substitution, σ, or a higher intensity of intermediates, $1/a$, imply higher North-South unemployment differentials. As it is well known from the new economic geography literature, these parameters imply stronger agglomeration forces and hence wider disparities once the symmetric equilibrium is broken. A higher ε implies lower migration costs and hence higher mobility. For very high values of ε, migration costs are not high enough to impede (almost) complete agglomeration of manufacturing firms when the symmetric equilibrium becomes unstable. We are not interested in this possibility, because it implies that the problem of peripheral unemployment disappears, since almost no manufacturing workers would be left in the periphery.[15]

The labor market parameters are s, the rate of job destruction, β, the share of the match surplus that goes to workers and η, the elasticity of the matching function to the unemployment rate. Both s and β imply higher regional unemployment rates, but have a small impact on regional disparities. On the other hand, a higher elasticity η induces higher unemployment, but lower regional disparities, since it deteriorates the labor market conditions less than proportionately in the region with a higher unemployment rate.

Finally, a higher lump sum tax T allows the government to finance higher unemployment benefits. This deteriorates the regional labor markets and induces higher unemployment rates in both regions. Since the subsidy is fixed in nominal terms and geographically undifferentiated, it generates a stronger distortion where prices are lower (because of its higher real value). Therefore, a higher subsidy improves the relative performance of the Southern labor market as symmetry is broken. This effect vanishes as trade integration generates price converge.

3.2 Empirical implications

Figure 5.2 summarizes the main implications of the model. The horizontal axis reports the level of trade costs. The solid line represents the North-South unemployment gap (in percentage) as a function of regional integration. The dashed line represents the North-South gap (in percentage) in terms of real per capita income. Finally, the dotted line represents the percentage of immigrants in Northern population. The figure tells the following story. Starting from a symmetric equilibrium, a gradual regional integration triggers migrations to the core and this leads to the emergence of strong regional disparities, both in terms of per capita income and un-

[15] Note that, in Figure 1, when the symmetric equilibrium loses stability, it gives rise to two stable steady states in its neighborhood. This kind of bifurcation is called "supercritical pitchfork" (see Ottaviano [2000] for an illustration of nonlinearities arising in new economic geography models, and Puga [1999]). For a sufficiently high ϵ, the number and stability of steady states changes. In this latter case, for trade costs in the neighorhood of the level at which the symmetric equilibrium becomes unstable, two unstable interior equilibria appear around the symmetric equilibrium. When these equilibria disappear, they give the symmetric equilibrium their instability. This alternative bifurcation is called "subcritical pitchfork". Further, in the former case the evolution of steady-states as trade costs are reduced is gradual, whereas in the latter there is a discontinuous change.

employment. Thereafter, further falls in trade costs bring about convergence in per capita income. The reason for this result is that the volume of imports is relatively higher in the periphery, and hence it gains relatively more than the core in terms of real income.

Note that, during the phase of convergence in per capita income, the core-periphery unemployment gap grows larger until a higher level of integration is reached. This happens because the sharp price fall in the periphery raises the real value of unemployment subsidies, and further deteriorates the labor market conditions in that region.

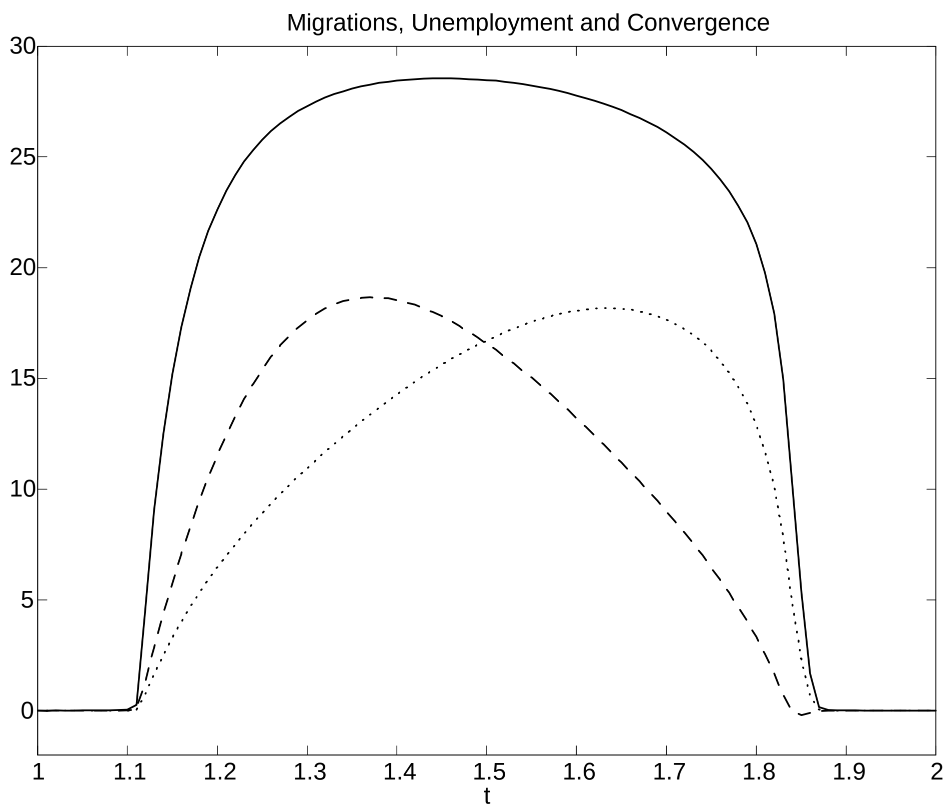

Figure 5.2. market integration and regional disparities

Finally, in this economy migrations happen in waves. The model predicts large migrations when symmetry is broken and similarly large return migrations shortly before complete convergence, but very little migrations in between. Therefore, the percentage of immigrants in the North goes up sharply when agglomeration starts, stays almost constant for a substantial range of trade costs and then declines when symmetry is gradually restored.

The model provides a stylized conceptual framework which can help explain the empirical puzzles mentioned in the introduction. In fact, if we assume that in the early Eighties European regions were already in a core-periphery equilibrium, then the model predicts convergence in real per capita income, divergence in regional unemployment rates, and declining migration rates after regional integration, i.e., it can explain the most striking features of the last decades European regions' evolutions.

4 Transitional dynamics

In this section we explore the adjustment path which leads the system from one steady-state to another after a fall in trade costs. In particular, we analyze the transitional dynamics of the system after a once and for all unanticipated fall in transport costs. In order to accomplish this we have linearized around the steady-state the system described by equations (5.3), (5.4), (5.7)-(5.16), and by the equivalent equations for region j[16]. We have chosen as a point of departure for our analysis an equilibrium in which manufacturing is already partially agglomerated in the North. The reason is that our purpose is to explain some stylized facts concerning the regional evolution of unemployment in Europe during the last decades, when regional disparities were already pronounced. The results are shown in Figure 5.3. The graphs plot the adjustment path after a 10% fall in trade costs, from $\tau = 1.5$ to $\tau = 1.45$. The model is calibrated for quarterly data, therefore each period corresponds to three months.

Note that the dynamic system which governs the short run adjustment has only two state variables, namely the employment levels in the two regions. As the matching process between jobs and workers requires time, the response of employment levels to a change in the environment is gradual. No other variable is assumed to be sluggish.

[16] The choice of a local solution method is dictated by computational tractability. Even though the original system is non linear, our approximation can be considered reliable because we only study the dynamic adjustment between steady-states which are fairly close to each other. Further, our main interest is on the qualitative behavior of the model rather than on quantitative predictions.

Panel (a) plots the time path of the total manufacturing work force and the employment level in the South. In this exercise, the reduction of trade costs reinforces the geographical advantage of the North, which makes the core region more attractive for locating manufacturing firms and workers. The result is a wave of migration from the periphery and a discrete jump in regional labor force. As already noted, the reaction of employment is gradual: it falls smoothly in the South, because the rate of job destruction is not compensated any more by new matches. Symmetrically, it rises gradually in the North, where the higher number of job seekers increases the likelihood of a match. The eventual increase in employment in the North and the fall in the South strengthens even more the agglomeration forces in the core region. This implies further migration flows from the periphery (although at a slower pace), and further agglomeration of production in the core, until the new steady state is reached.

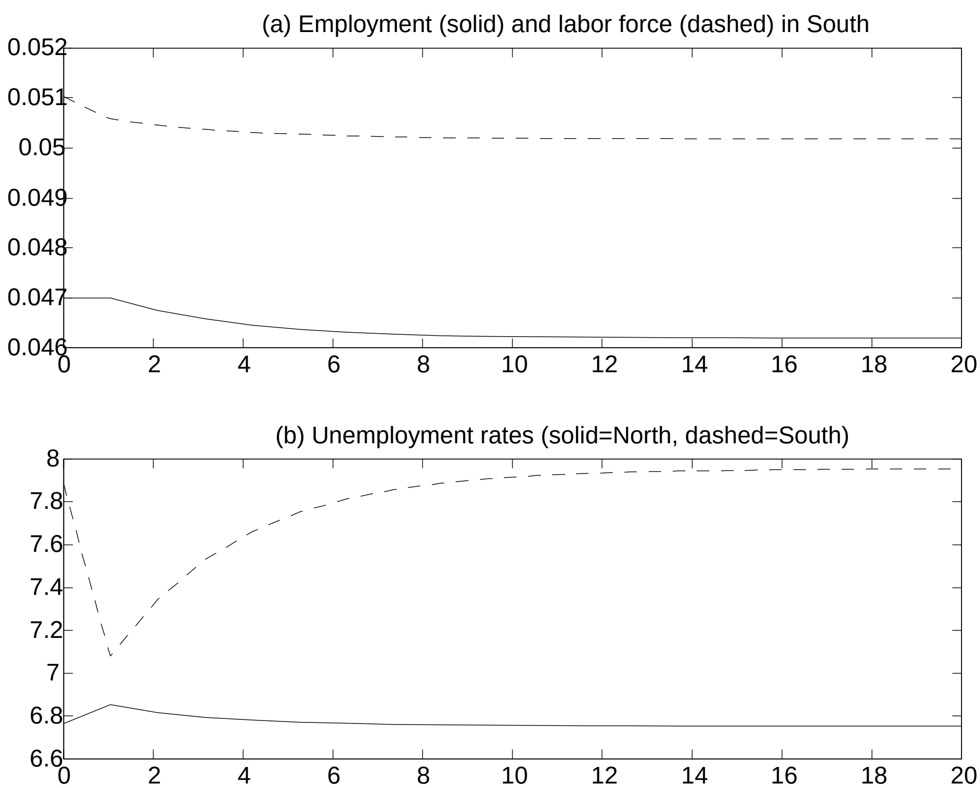

Figure 5.3. dynamic adjustment after a trade shock

Panel (b) shows the evolution of regional unemployment rates. As unemployed workers move from the South to the North, the impact effect of a fall in trade costs is a temporary discrete fall in the unemployment rate of the South and a rise in the North. As manufacturing production agglomerates in the core, the unemployed workers are gradually absorbed; moreover, the consequent fall in the price of intermediates reduces the search cost for Northern firms and this improves the labor market conditions. The opposite happens in the South. Therefore, after the first jump, the unemployment rates in the two regions diverge.

Two points are worth noting. First, during the transition migration is gradual. We obtain this result without imposing any ad hoc sluggishness on labor mobility. The reason for our result is that as incoming migrants are gradually employed, the geographical advantage of the North is reinforced and this attracts more workers from the South.

Second, and more importantly, the transitional dynamics highlight a trade-off between the short run and long run effects of migration flows on the core-periphery unemployment gap: in the presence of strong agglomeration forces and inefficiencies in the job matching process, the migration flows induced by regional integration cause a temporary convergence in the regional rates of unemployment. This happens because migrations reduce the pool of unemployed workers in the South and expand it in the North. However, this induces a positive externality on searching firms in the North and a negative externality on Southern firms. The result is an increase in employment in the North and a fall in the South which strengthens agglomeration forces in the North and reduces them in the South. Hence, in the new steady-state, when Southern immigrants are absorbed by the Northern labor market, the North-South unemployment gap is permanently higher than before the trade shock.

How long does it take for the system to adjust after a trade shock? Simulations reported in Figure 5.3 are drawn for the same parameters values used in Figures 5.1-2, and are calibrated for a time unit of one quarter. The time path of the variables shows that the adjustment is fairly fast, as the transition after a large trade shock is almost complete in less then three years. An interesting implication of a fast transition is that it allows us to give a broader interpretation of the simulation results reported in the preceding section. Although the graphs in Figures 5.1-2 are simply a collection of steady-states as a function of trade costs, we can interpret them as an approximation of the evolution of regional variables after a sequence of steps along a process of regional integration. Further, with reference to the evolution of regional unemployment, the analysis of the transition suggests a picture even more extreme than the one reported in panel (c) of Figure 5.1. As long as migration rates to the North remain positive, the short run dynamics will tend to reduce the unemployment gap, so that large unemployment differentials will become evident when migration rates are almost nil.

5 A policy experiment

The simple model developed so far does not provide a suitable framework for policy evaluations; the presence of various dimensions of heterogeneity (workers differ according to geographical location, employment status and sector of production) leads to difficulties in defining an aggregate welfare function. However the model, as simple as it is, includes a policy variable, the level of the subsidy, which captures an important feature of the European welfare system. A large body of literature has blamed the excessive level of subsidization of European countries for the high rates of unemployment experienced in the past decades and a growing number of studies has used similar arguments to question the sustainability of the welfare state (see, among others, Ljungqvist and Sargent [1998]). It is therefore natural to ask what is the effect of a reduction in the generosity of the common unemployment policy on a geographically differentiated economy.

Figure 5.4 shows the dynamic adjustment after a once and far all unanticipated reduction in the lump sum tax, from $T = 0.024$ to $T = 0.02$. This implies a 9% decrease of the nominal subsidy. Trade costs are set at the intermediate level $\tau = 1.5$, which corresponds to an asymmetric equilibrium. It is apparent that the long run effect is a reduction of the unemployment rates, both in the core and in the periphery. But the effect is generally uneven among the two regions.

Panel (a) shows that a cut in the subsidy strengthens the geographical advantage of the North and therefore induces migrations from the periphery. This happens because the subsidy, set in nominal terms, generates a higher distortion in the region with a lower price index (the core), hence the North benefits relatively more from the cut.[17]

Panel (b) illustrates again a tension between the short run and long run adjustment: as job seekers leave the South, the unemployment gap is initially reduced, but the gradual strengthening of agglomeration economies in the North leads to an improvement in the relative performance of the Northern labor market. However, the picture suggests that the costs of the transition will be borne more than proportionally by the core region.

Figure 5.4. dynamic adjustment after a policy shock

[17]Note that the price index q_i is not to be confused with empirical price indices, such as the CPI. The reason is that q_i is a perfect price index which fully reflects the regional availability of consumption and intermediate goods. Given the assumption of love for variety embedded both in the preferences and the production functions, it follows that the price of utility is higher in the peripheral regions.

6 Final remarks

In the last decades Western Europe has undergone a process of deep economic integration. Recent developments in the field of the new economic geography have shown that such a process may trigger the spatial agglomeration of economic activity. However, this literature neglects any imperfections in the labor market and hence it cannot explain the geography of unemployment. Yet, the evidence concerning European regions shows a strong tendency toward polarization and divergence of regional unemployment rates, together with a slight tendency toward convergence in per capita income. As a consequence, the uneven spatial distribution of unemployment is nowadays the main cause of policy concern in Europe.

In this paper, we have formulated a core-periphery model with frictions in the job matching process, in order to study the coevolution of income, migrations and unemployment rates at regional level. We have shown how market integration can be a driving force behind a recently documented empirical puzzle: the divergence of unemployment rates, together with low mobility and modest income convergence experienced by European regions over the last twenty years.

By studying explicitly the transitional dynamics of the model we have also highlighted a contrast between short run and long run effects of a shock on a geographically differentiated economy. In particular, our model illustrates how labor mobility can alleviate temporarily regional disparities but it exacerbates them in the final adjustment. This tension between short run and long run responses may shed some light on the mixed evidence concerning the labor market effects of immigration (see, for instance, Borjas [1999] and Borjas et al.[1997]).

Since the main interest of this paper was on the qualitative behavior of regional macro variables, we adopted a very stylized framework which enabled us to keep a high level of generality. To address quantitative questions and to match more closely empirical data, the structure of the model could be made more realistic by introducing asymmetries in the underlying economic structure of the regions.

Our model is too simple to lend itself to any robust policy prescription. However, if we take seriously some of its logical implications, it would suggest the policy maker to do any effort to further reduce the core-periphery transportation costs, since unemployment differentials tend to disappear when these costs become negligible. Practically, this may require a strengthening of communication networks (e.g., road and telecommunication infrastructure) in order to facilitate access to larger markets for peripheral regions. In this respect, Viesti (2000) reports some supportive empirical evidence: he shows that 25 industrial clusters in Southern Italy have almost nothing in common, but the proximity to highways.

However, we have also shown that such a process of falling distance costs may further increase the core-periphery unemployment gap before the process of convergence definitely sets in. Hence, we may still observe for some time a further deterioration of labor market conditions in European peripheral regions.

7 References

Barro, Robert and Xavier Sala-i-Martin (1991). "Convergence across States and Regions", *Brookings Papers on Economic Activity* 1, 107-182.

Bentolila, Samuel (1997). "Sticky Labor in Spanish Regions", *European Economic Review* 41, 591-598.

Borjas, George J. (1999). "The economic analysis of immigration" in O. Ashenfelter and D Card (eds), *Handbook of Labor Economics*, Vol. 3A, North-Holland.

Borjas, George J., Richard B. Freeman and Lawrence F. Katz (1997). "How much do immigration and trade affect labor market outcomes", *Brooking Papers on Economic Activity*, 1-67.

Davidson, Carl, Lawrence Martin and Steven Matusz (1999). "Trade and Search Generated Unemployment", *Journal of International Economics* 48, 271-299.

Faini, Riccardo (1996). "Increasing Returns, Migration and Convergence," *Journal of Development Economics* 49, 121-136.

Faini, Riccardo (1997), Gianpaolo Galli, Pietro Gennari and Fulvio Rossi. "An empirical Puzzle: Falling Migration and Growing Unemployment Differentials among Italian Regions", *European Economic Review* 41, 571-579.

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

Jansen, Marion and Alessandro Turrini (1998). "Job Creation, Job Destruction and the International Division of Labour", *Centro Studi Luca d'Agliano D.S. Working Paper* n. 122.

Krugman, Paul (1991). "Increasing Returns and Economic Geography," *Journal of Political Economy* 99, 483-499.

Ljungqvist, Lars and Thomas J. Sargent (1998). "The European Unemployment Dilemma", *Journal of Political Economy* 106, 514-550.

Marshall, Alfred (1920). *Principles of Economics*, London, Macmillian.

Obstfeld, Maurice and Giovanni Peri (1998). "Regional Non-Adjustment and Fiscal Policy", *Economic Policy*, 207-259.

Ottaviano, Gianmarco and Diego Puga (1999). "Agglomeration in the Global Economy: A Survey of the New Economic Geography", *The World Economy* 21, 707-731.

Ottaviano, Gianmarco (2000). "*Ad Usum Delphini*: A Primer in 'New Economic Geography"', *Giornale degli Economisti e Annali di Economia* 59, 89-116.

Pissarides, Christopher (1985). "Short-Run Equilibrium Dynamics of Unemployment, Vacancies, and Real Wages," *American Economic Review* 75, 676-690.

Pissarides, Christopher (1990). *Equilibrium Unemployment Theory*, Basil Blackwell, Oxford.

Pissarides, Christopher (1998). "The Impact of Employment Tax Cuts on Unemployment and Wages; The role of Unemployment Benefits and Tax Structure", *European Economic Review* 42, 155-183.

Puga, Diego (1999). "The Rise and Fall of Regional Inequalities", *European Economic Review* 43, 303-334.

Puga, Diego and Henry Overman (1999). "Unemployment Clusters across European Regions and Countries", *CEPR Discussion Paper* n. 2255.

Quah, Danny (1996). "Regional Convergence Clusters across Europe", *European Economic Review* 40, 951-958.

Quah, Danny (1997). "Increasingly Weightless Economies", *Bank of England Quarterly Bulletin* 37(1), 49-56.

Viesti, Gianfranco (2000). *Come Nascono i Distretti Industriali*, Editori Laterza, Bari.

8 Appendix

Unless otherwise stated in the text, all the graphs are drawn using the following parameter values: $\sigma = 5$, $\mu = 0.5$, $\alpha = 0.25$, $\beta = 0.5$, $r = 0.02$, $\eta = 0.4$, $s = 0.18$, $a = 20$, $\epsilon = 11$, $T = 0.024$, $d = 0.025$.

MONOGRAPH SERIES

1. Michaely, Michael *The Theory of Commercial Policy: Trade and Protection,* 1973
2. Söderström, Hans Tson *Studies in the Microdynamics of Production and Productivity Change*, 1974
3. Hamilton, Carl B. *Project Analysis in the Rural Sector with Special Reference to the Evaluation of Labour Cost,* 1974
4. Nyberg, Lars and Staffan Viotti *A Control Systems Approach to Macroeconomic Theory and Policy in the Open Economy*, 1975
5. Myhrman, Johan *Monetary Policy in Open Economies*, 1975
6. Krauss, Melvyn *International Trade and Economic Welfare*, 1975
7. Wihlborg, Clas *Capital Market Integration and Monetary Policy under Different Exchange Rate Regimes*, 1976
8. Svensson, Lars E.O. *On Competitive Markets and Intertemporal Resources Allocation*, 1976
9. Yeats, Alexander J. *Trade Barriers Facing Developing Countries*, 1978
10. Calmfors, Lars *Prices, Wages and Employment in the Open Economy*, 1978
11. Kornai, János *Economics of Shortage*, Vols I and II, 1979
12. Flam, Harry *Growth, Allocation and Trade in Sweden. An Empirical Application of the Heckscher-Ohlin Theory*, 1981
13. Persson, Torsten *Studies of Alternative Exchange Rate Systems. An Intertemporal General Equilibrium Approach*, 1982
14. Erzan, Refik *Turkey's Comparative Advantage, Production and Trade Patterns in Manufactures. An Application of the Factor Proportions Hypothesis with Some Qualifications*, 1983
15. Horn af Rantzien, Henrik *Imperfect Competition in Models of Wage Formation and International Trade*, 1983
16. Nandakumar, Parameswar *Macroeconomic Effects of Supply Side Policies and Disturbances in Open Economies*, 1985
17. Sellin, Peter *Asset Pricing and Portfolio Choice with International Investment Barriers*, 1990
18. Werner, Ingrid *International Capital Markets: Controls, Taxes and Resources Allocation*, 1990
19. Svedberg, Peter *Poverty and Undernutrition in Sub-Saharan Africa: Theory, Evidence, Policy*, 1991
20. Nordström, Håkan *Studies in Trade Policy and Economic Growth*, 1992
21. Hassler, John, Lundvik, Petter, Persson, Torsten and Söderlind, Paul *The Swedish Business Cycle: Stylized facts over 130 years*, 1992
22. Lundvik, Petter *Business Cycles and Growth*, 1992
23. Söderlind, Paul *Essays in Exchange Rates, Business Cycles and Growth*, 1993
24. Hassler, John A.A. *Effects of Variations in Risk on Demand and Measures of Business Cycle Comovements*, 1994
25. Daltung, Sonja *Risk, Efficiency, and Regulation of Banks*, 1994
26. Lindberg, Hans *Exchange Rates: Target Zones, Interventions and Regime Collapses*, 1994
27. Stennek, Johan *Essays on Information-Processing and Competition*, 1994
28. Jonsson, Gunnar *Institutions and Incentives in Monetary and Fiscal Policy*, 1995
29. Dahlquist, Magnus *Essays on the Term Structure of Interest Rates and Monetary Policy*, 1995
30. Svensson, Jakob *Political Economy and Macroeconomics: On Foreign Aid and Development,* 1996
31. Blix, Mårten *Rational Expectations and Regime Shifts in Macroeconometrics,* 1997
32. Lagerlöf, Nils-Petter *Intergenerational Transfers and Altruism,* 1997
33. Klein, Paul *Papers on the Macroeconomics of Fiscal Policy,* 1997
34. Jonsson, Magnus *Studies in Business Cycles,* 1997
35. Persson, Lars *Asset Ownership in Imperfectly Competitive Markets,* 1998
36. Persson, Joakim *Essays on Economic Growth*, 1998
37. Domeij, David *Essays on Optimal Taxation and Indeterminacy,* 1998
38. Flodén, Martin *Essays on Dynamic Macroeconomics,* 1999
39. Tangerås, Thomas *Essays in Economics and Politics: Regulation, Elections and International Conflict*, 2000
40. Lidbom, Per Pettersson *Elections, Party Politics and Economic Policy*, 2000

41. **Vestin, David** ***Essays on Monetary Policy,*** **2001**
42. **Olofsgård, Anders** ***Essays on Interregional and International Political Economics,*** **2001**
43. **Johansson, Åsa** ***Essays on Macroeconomic Fluctuations and Nominal Wage Rigidity,*** **2002**
44. **Groth, Charlotta** ***Topics on Monetary Policy,*** **2002**
45. **Gancia, Gino A.** ***Essays on Growth, Trade and Inequality,*** **2003**