Essays on Economic Growth and the Skill Bias of Technology

by

Nico Voigtländer

Department of Economics
Universitat Pompeu Fabra
c/Ramon Trias Fargas 25-27
08005 Barcelona, Spain
Email: nico_v@alum.mit.edu
Phone: +34-67 062 3218

Submitted to the Department of Economics on March 20, 2008, in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Economics

Abstract

My dissertation is a collection of three essays that consider various aspects of long-run economic growth as well as income inequality and the demand for skilled labor.

The first chapter, co-authored with Joachim Voth, investigates the questions why England was the first country to industrialize. We present a probabilistic two-sector model where the initial escape from Malthusian constraints depends on the demographic regime, capital deepening and the use of more differentiated capital equipment. Weather-induced shocks to agricultural productivity interact with the demographic regime and affect the speed of growth. We calibrate our model to match the main characteristics of the English economy in 1700 and the observed transition until 1850. Higher initial per capita incomes together with fertility limitation increase industrialization probabilities. In contrast to unified growth theory in the tradition of Galor-Weil (2000) and Galor-Moav (2002), our setup does not depend on human capital accumulation and is therefore closer to the empirics. Since there is little evidence that human capital increased before 1850, this solves an important shortcoming in the existing literature. Simulations using parameter values for other countries show that Britain's early escape was only partly due to chance. France could have moved out of agriculture and into manufacturing faster than Britain, but the probability was less than 30 percent. Contrary to recent claims in the literature, 18th century China had only a minimal chance to escape from Malthusian constraints. [This chapter is published in: *Journal of Economic Growth* 2006, 11(4): 319-361.]

The second chapter is motivated by the finding that relatively high initial incomes in 1700 gave European countries the edge to industrialize. This chapter is also co-authored with Joachim Voth. Using a simple Malthusian model with two sectors, we examine why Western Europe overtook China in terms of incomes and urbanization rates in the early modern period (1450-1700). Standard accounts of this "reversal of fortune" emphasize European inventiveness and the slackening of Chinese technological creativity (Mokyr, 1990). That living standards could exceed subsistence levels in a Malthusian setting at all should be surprising. Rising fertility and falling mortality ought to have reversed any gains. We show that productivity growth in Europe can only explain a tiny fraction of rising living standards. Population dynamics – changes of the birth and death schedules – were far more important drivers of the long-run Malthusian equilibrium. In our setup, population fell following the Black Death; wages surged. Because of Engel's Law, demand for urban products increased, raising urban wages and attracting rural population. European cities were particularly unhealthy; urbanization pushed up aggregate death rates. This effect was reinforced by more frequent wars, fed by city wealth, and disease spread by trade. Thus, higher wages themselves reduced population pressure. Without technological change, our model can account for income increases that led to levels far above subsistence, as well as the sharp rise in European urbanization.

Human capital accumulation is at the heart of unified growth theory. The transition from stagnation to growth in these models goes hand in hand with an increasing importance of skills. Historical observations suggest the opposite: The first stage of the Industrial Revolution was skill-replacing rather than skill-using. It was only later on that technical change became skill biased. Yet, we do not fully understand what caused this change in the nature of technologically-induced factor demand. Previously suggested explanations like international trade or complementarities between technology and skills cannot account for the sheer magnitude of the observed skill bias in recent decades. This motivates the third chapter. I present a novel stylized fact and analyze its contribution to the skill bias of technical change: The share of skilled labor embedded in intermediate inputs correlates strongly with the skill share employed in final production. This finding points towards an intersectoral technology-skill complementarity (ITSC). Empirical evidence suggests that the channel through which this complementarity works is product innovation driven by skilled workers. Together with input-output linkages, the observed complementarity delivers a multiplier that reinforces skill demand along the production chain. The effect is large, accounting for more than one third of the observed skill upgrading in U.S. manufacturing over the period 1967-92. I also present a simple multi-sector model with intermediate linkages that integrates the observed ITSC into the standard framework of skill-biased technical change. Therein, the relative productivity of skilled workers rises with the skill intensity of intermediates. A calibration exercise confirms the quantitative importance of the ITSC.

Thesis Supervisor: Hans-Joachim Voth
Title: ICREA Research Professor, Universitat Pompeu Fabra

Acknowledgements

First and foremost, I thank my advisor Joachim Voth. His helpful suggestions, comments, and constructive criticism have been invaluable. The lessons he taught me greatly influenced the way I think about economics. His guidance and support went far beyond the usual commitment. Over the past years, Joachim has not only been a magnificent advisor and brilliant co-author, but also a great friend.

I also thank Antonio Ciccone and Jaume Ventura for their advice throughout my thesis writing. Paula Bustos, Fernando Broner, Gino Gancia, and Diego Puga also deserve thanks for their suggestions and encouragement, especially during my research on the jobmarket paper.

Many of my fellow students at UPF have contributed in important ways to this project. In particular, I thank Francesco Turino for his friendship, encouragement, helpful suggestions, and entertainment beyond economics. I have also greatly benefited from discussions with Alexander Ludwig who came to UPF as a visiting student in 2004 and has been my friend ever since.

This thesis has also benefited from my visit to UC Berkeley during the last year of my PhD. I am grateful to Christina Romer for making this visit possible. I also thank Yuriy Gorodnichenko for his creative suggestions and the frequent discussion of my research.

One person is connected like no other to everyday life at UPF economics: The secretary Marta Araque. She is what one might call the soul of the graduate program. Throughout my studies at UPF, her door was always open, she never seemed to need a break, nor could huge piles of work strike her with awe. I am deeply grateful for her support and her guidance around the obstacles of administration.

Finally and most importantly, I thank my girlfriend Gabriela and my family for their support and encouragement. I am grateful to my mother, my brother and sister, and especially my grandparents for making home the most kind-hearted place in the world.

Dipòsit legal: B.3615-2009
ISBN:

Table of Contents

1 Why England? Demographic Factors, Structural Change and Physical Capital Accumulation during the Industrial Revolution **7**

1.1 Introduction . . . 7
1.2 Historical background and motivation . . . 9
1.3 The Model . . . 12
1.3.1 Consumers . . . 12
1.3.2 Production . . . 14
1.3.3 Final goods . . . 14
1.3.4 Capital . . . 14
1.3.5 Model Representation with Aggregate Externalities . . . 15
1.3.6 Equilibrium and Industrialization . . . 16
1.3.7 Equilibrium and Industrialization without Population Dynamics . . . 16
1.3.8 Open Economy Considerations . . . 17
1.3.9 Inequality . . . 18
1.3.10 Population Dynamics . . . 19
1.4 Calibration and Simulation Results . . . 22
1.4.1 Calibration . . . 22
1.4.2 The Industrial Revolution in England . . . 25
1.4.3 Sensitivity Analysis . . . 26
1.4.4 The Role of Chance . . . 28
1.4.5 Probabilities of Industrialization in other Countries . . . 29
1.4.6 Turning England into China . . . 32
1.5 Conclusions . . . 33
1.6 APPENDIX . . . 35
1.6.1 Appendix A.1 – Optimization of Production . . . 35
1.6.2 Appendix A.2 – Capital Varieties and Aggregate Capital . . . 37
1.6.3 Appendix A.3 – Market Clearing and Equilibrium . . . 39
1.6.4 Appendix A.4 – Aggregate Externality Representation . . . 40
1.6.5 Appendix A.5 – Equilibrium Conditions of the 2-Sector Model . . . 41
1.6.6 Appendix A.6 – Numerical Simulations . . . 42
1.6.7 Appendix A.7 – Calibration of the British Birth Function . . . 43
References . . . 45

2 The Three Horsemen of Growth: Plague, War and Urbanization in Early Modern Europe **50**

2.1 Introduction . . . 50
2.2 Historical context and background . . . 53
2.3 The Model . . . 57
2.3.1 Consumption . . . 58
2.3.2 Production . . . 59
2.3.3 Migration . . . 59

2.3.4 Population Dynamics . . . 61
2.3.5 Equilibria . . . 62
2.4 Calibration and Simulation Results . . . 67
2.4.1 Calibration . . . 67
2.4.2 Plague and Equilibrium without Horsemen Effect . . . 69
2.4.3 Long-run Equilibrium with Horsemen Effect . . . 70
2.4.4 Technological Progress and Model Fit . . . 71
2.5 Extensions . . . 71
2.5.1 The European Marriage Pattern . . . 71
2.6 Concluding Remarks . . . 73
References . . . 75

3 Many Sectors Meet More Skills: Intersectoral Linkages and the Skill Bias of Technology 79
3.1 Introduction . . . 79
3.2 Motivation and Framework . . . 81
3.3 Data . . . 87
3.4 Empirical Results . . . 91
3.4.1 Correlation of Skill Intensities across Sectors . . . 92
3.4.2 Robustness of the Correlation . . . 94
3.4.3 Investigating the Channel of the ITSC . . . 98
3.4.4 The ITSC's Contribution to Skill Upgrading . . . 102
3.5 A Sketch Model . . . 108
3.5.1 Production and Consumption . . . 109
3.5.2 Linkages, Complementarities, and Multipliers . . . 110
3.5.3 Optimization and the Symmetric Case . . . 110
3.5.4 Calibration . . . 112
3.6 Conclusions . . . 114
3.7 APPENDIX . . . 116
3.7.1 Appendix A.1 Data Sources and Construction of Variables . . . 116
3.7.2 Appendix A.2 Equilibrium for the Symmetric Case . . . 117
3.7.3 Appendix A.3 The Multiplier in the Model with N Sectors . . . 119
References . . . 121

Chapter 1

Why England? Demographic Factors, Structural Change and Physical Capital Accumulation during the Industrial Revolution

(Joint with Joachim Voth, UPF)

1.1 Introduction

Britain was the first country to break free from Malthusian constraints, with population size and living standards starting to grow in tandem after 1750 [Crafts (1985), Wrigley (1983)]. In many parts of the world, however, growth rates of per capita income took a long time to accelerate. Eventually, more and more countries industrialized, first in Europe and North America, and from the 20th century onwards in other areas of the globe. The relative size of economies, the onset of the demographic transition, and living standards of citizens are still profoundly influenced by the timing of Industrial Revolutions around the globe [Galor and Mountford (2003)] – with dramatic consequences for the economic and political history of the world that are still felt today.

Why did some countries industrialize so much earlier than others? Unified growth theory [Galor and Weil (2000), Galor and Moav (2002), Jones (2001), Hansen and Prescott (2002)] offers a consistent explanation for the transition from century-long Malthusian stagnation to rapid growth. What is missing is a better understanding of why some countries overcame stagnation at radically different points in time. The question is almost as old as industrialization itself. Economic historians have stressed a long list of factors, ranging from the property rights regime to the land tenure system, that might have favored Britain [Landes (1999)]. Galor (2005) argues that geographical factors and historical accident interacted to delay or accelerate the timing of the "Great Escape", and that "variations in institutional, demographic, and cultural factors, trade patterns, colonial status, and public policy" may have played a role. This paper aims to provide a systematic answer to the questions "Why England?" and "Why Europe?" In doing so, it offers clear quantitative evidence on the role of starting conditions and the nature of constraints that delayed industrialization for centuries in many parts of the world.

In our model, chance can play an important role. Industrialization is treated as the result of a probabilistic process. During the late medieval and early modern period, brief expansions – "efflorescences" – occurred in many countries [Braudel (1973), Goldstone (2002)]. Yet most of these growth episodes sooner or later ground to a halt. Some advanced economies (such as the Italian Republics) went into decline, while countries like the Netherlands stagnated at high income levels. This is why economic historians have often been sceptical of industrialization theories where the final outcome is pre-determined [Clark (2003a), Mokyr and Voth (2007)]. What explains these starts and stops? And could other countries have succeeded before Britain? Crafts (1977) argued that accidental factors, and not systematic advantages, may have been crucial – that France, for example, could have easily industrialized first had it not been for a number of random factors. To examine the determinants of early economic development, this paper develops a simple stochastic model of the first Industrial Revolution – the transition from the Malthusian to the post-Malthusian regime, in the terminology of unified growth theory. In the spirit of Stokey (2001), our model is then calibrated with eighteenth-century English data. We find that chance played a role in the timing and speed of Britain's initial surge – it's actual performance was at the upper half of the expected range of outcomes in our model. By altering the parameters of the calibrated model, we derive probabilities of the escape in other parts of the world. France could have experienced substantial growth, based on our model, but the manufacturing employment share in 1850 is lower than in Britain in most of our simulations.

As emphasized by Galor and Moav (2004), physical capital accumulation is crucial for the first transition. This is reflected in our model, which emphasizes TFP advances as a result of growing capital inputs. The key factors influencing industrialization probabilities in our model are starting incomes, the nature of shocks, inequality, and the demographic regime. In our calibrations, we find that England's (and Europe's) chances of sustained growth were greater principally because the demographic regime propped up initial incomes. Redistribution plays only a small role. Galor and Moav (2004) argue that inequality should be beneficial for industrialization in its initial stages, when physical capital is crucial; during the second transition to self-sustaining growth, human capital becomes a key input, and inequality is harmful. Zweimüller (2000) shows how, in an endogenous growth model, redistribution can be growth-enhancing, while Matsuyama (2002) demonstrates how development depends on the exact shape of the income distribution. We add another dimension emphasized by Fogel (1994). As many as 20 percent of the population in 18th century France possibly did not receive enough food to work for more than a few hours a day. Also, when inequality was too great prior to the Industrial Revolution, crisis mortality could be high. This undermines growth by lowering the marginal return to capital, and the pace of accumulation. If this effect is larger than the rise in the capital/labor and land/labor ratios due to falling population, productivity growth suffers. We conclude that inequality may only be beneficial via the savings channel if the population is sufficiently well-fed to avoid famines and chronic undernutrition.

Our work is related to three bodies of literature. Economic historians have sometimes been sceptical of endogenous growth models.[1] Crafts (1995) rejected endogenous growth models partly because they had little to say about the different speeds of industrialization in England and France. He also argued that detailed accounts of technological historians did not square with the predictions of endogeneous growth models. Unified growth theory has made considerable progress in bridging the gap between theory and historical facts. We therefore take the unified growth models by Galor and Moav (2002), Galor and Weil (2000), Jones (2001), Kögel and Prskawetz (2001) and Cervellati and Sunde (2005) as our starting point. Our model focuses on what Galor et al. call the first of two crucial transitions – the one from Malthusian to a post-Malthusian world, when population pressure no longer determines wages (but before human capital becomes crucial). In the vein of these models, demographic feedback and physical capital accumulation are important for the initial escape from stagnation. While papers in the Galor-Weil tradition focus on fertility limitation after the first transition, we emphasize the importance of fertility behavior for starting conditions in Europe (as in the work of Wrigley (1988), inter alia).

A second set of related papers emphasizes technology adoption. Murphy, Shleifer and Vishny (1989a,b) argue that bigger markets and moderate inequality facilitate the adoption of new technologies when fixed

[1] Voth (2003) concluded that "the Industrial Revolution in most growth models shares few similarities with the economic events unfolding in England in the 18th century".

costs are substantial. The technological history of the First Industrial Revolution only offers qualified support for the importance of fixed costs and indivisibilities. Instead, we employ an externality to capital use that is based on the findings of technological historians (Mokyr 1990). As in Kögel and Prskawetz (2001), we emphasize the interactions between agricultural productivity and industrial growth – an approach that goes back in economic history to Gilboy (1932). Acemoglu and Zilibotti (1997) observe that volatility in poor economies is high. New technologies represent high risk, high return investment. Because of indivisibilities, only richer and larger countries undertake them. A run of "good years" increases the probability of switching to high-productivity projects. In our model, stochastic income fluctuations and starting per capita income play a role because they increase the scope for the capital externality to work.

The third body of literature uses calibrations and simulation methods to shed new light on the industrialization process. Stokey (2001) was amongst the first to employ calibrations for the Industrial Revolution. She concludes that foreign trade and technological change in manufacturing were equally important for growth, but that improvements in energy production mattered less. Crafts and Harley (2000) examine the importance of broad-based technological change in a CGE model, and conclude that slow, sector-specific improvements in TFP are compatible with the observed pattern of trade. Lagerlöf (2003) uses a probabilistic model where mortality fluctuations – epidemics – eventually lead to a transition to self-sustaining growth. Lagerlöf (2006) simulates the Galor-Weil model, and finds that it can replicate most of the important features in the transition from stagnation to growth. Our approach differs from the Stokey approach in that it uses a more explicit model of productivity change. We combine the calibration exercise with the probabilistic models in the spirit of Lagerlöf (2003).

The paper proceeds as follows. Section II discusses the historical context and motivation for the paper. It briefly highlights where existing unified growth models are in conflict with the historical record, and sets out the basic elements of our story. Section III presents the model, explaining the role of demographic factors and the productivity benefits of differentiated capital inputs. In the next part, we calibrate the model and derive comparative industrialization probabilities for Britain, France, and China. Section V concludes.

1.2 Historical background and motivation

We focus on three features of the First Industrial Revolution – the slow, gradual nature of productivity growth and structural change, the role of inequality, and the nature of technological advances. Research in economic history over the last three decades has emphasized the slow and gradual nature of economic and structural change after 1750. Where once scholars argued for a few decades during which the transition to rapid growth occurred, a much more gradualist orthodoxy has taken hold [Crafts and Harley (1992), Antras and Voth (2001)]. As table 1.1 shows, total factor productivity growth rates were barely higher after 1750 than before. What is remarkable about the period after 1750 in Britain is not output growth or TFP performance as such, but the fact that accelerated population growth coincided with stagnant or slowly growing wages and output per head [Mokyr (1999)] – which makes the term “post-Malthusian” [Galor (2005)] particularly apt. During the period, and in line with unified growth theory, investment rates increased from about 7% of GDP in 1760 to 14% in 1840 [Crafts (1985)].

One implication of the gradualist school of thought is that per capita living standards in Britain must

Table 1.1: Output and Productivity Growth during the Industrial Revolution

(percent per annum)	Feinstein (1981)	Crafts (1985)	Crafts and Harley (1992)	Antras and Voth (2003)
Output				
1760-1800	1.1	1	1	
1801-1831	2.7	2	1.9	
1831-60	2.5	2.5	2.5	
Productivity				
1760-1800	0.2	0.2	0.1	0.27
1801-1831	1.3	0.7	0.35	0.54
1831-1860	0.8	1	0.8	0.33

have been quite high by 1750 already. This underlines the importance of starting conditions. One major factor was the nature of its demographic regime. As Wrigley and Schofield (1981) have argued, social and cultural norms limited fertility in early modern England in a way that few other societies did. This led to higher per capita incomes. England practiced an extreme form of the 'European marriage pattern' – West of a line from St. Petersburg to Triest, age at first marriage for women was determined by socioeconomic conditions, not age at first menarche [Hajnal (1965)]. This stabilized per capita living standards and avoided the waste of resources and human lives resulting from Malthus' 'positive' check, when population declines through widespread starvation. Within the European context, England was characterized by a low-pressure demographic regime – negative shocks to income were mainly absorbed by falls in fertility rather than increases in mortality [Wrigley and Schofield (1981); Wrigley et al. (1997)]. Both the higher level of per capita income produced by this demographic regime, and the way in which it was achieved, play a crucial role in our model.

Second, Britain was a highly unequal society in the 17th and 18th century [Lindert and Williamson (1982), Lindert (2000)]. Nonetheless, average British standards of consumption were relatively high compared to French ones, with a markedly higher minimum level of consumption. Fogel (1994) estimated that as a result of higher inequality and lower per capita output, the bottom 20-30% of the French population did not receive enough food to perform more than a few hours of work. This was partly a result of higher productivity overall – Fogel calculates that the British consumed some 17 percent more calories than their French counterparts. Yet the crucial factor may have been support for the poorer parts of society. The Old Poor Law was an unusually generous form of redistribution. At its peak, transfers amounted to 2.5% of British GDP, and more than 11% of the population received some form of relief. This may also have had indirect effects for the wages of those who were not recipients, by reducing competition in the labor market and raising the aggregate wage bill [Boyer (1990)]. Mokyr (2002) calculates that at its peak the system may have boosted average incomes of the bottom 40 percent of society by 14 to 25 percent. This ensured that in England, most individuals were sufficiently well-fed to work. It may have also stabilized consumer demand for industrial products. Even during the 1790s, when food prices were high, up to 30% of working class budgets continued to be spent on non-food items (with 6% going on clothing). With most of the goods produced by the nascent modern sector having high income elasticities of demand (in excess of 2.3), even modest gains in real wages in the later stages of the Industrial Revolution could translate into rapidly growing purchases of manufacturing goods. Finally, because of the large absolute value of the own-price elasticity of non-food spending (of –1.8 amongst the English poor), productivity increases and subsequent price reductions facilitated the growth of the modern sector [Horrell (1996)].

The third element in our story emphasizes the relative importance of innovation vs. inventions. Traditionally, economic historians in the tradition of North and Thomas (1973) have emphasized the importance of property rights, especially the patent system. In this view, as the security of property rights improved after the Glorious Revolution in 1688, more inventive activity took place. Technological change accelerated. The problem with this interpretation is that intellectual property rights were poorly protected before the 19th century in England, that few inventors received substantial material rewards, that the role of traditional ("feudal") forms of reward like grants from Parliament dominated benefits from patents, and that non-monetary incentives and chance seem to have played an extraordinarily large part in many of the key breakthroughs. Most of the technologies that made Britain great in 1850 were already known a century before. As Mokyr (1990) has emphasized, the crucial breakthroughs did not take the shape of blueprints or ideas. Instead, a stream of microinventions gave the First Industrial Nation its edge:

"In Britain, [...] the private sector on its own generated the technological breakthroughs and, more importantly, adapted and improved these breakthroughs through a continuous stream of small, anonymous microinventions which cumulatively accounted for the gains in productivity." [Mokyr (1990)]

New ways of adapting and making useful existing technologies were crucial. The Watt steam engine was but a variation of the Newcomen design. Many productivity advances were embodied in better pieces of capital equipment [Mokyr (1990)]. What made these advances possible was not a small group of heroic inventors but a small labor aristocracy of highly skilled craftsmen, perhaps no more than 5 percent of the workforce overall [Mokyr and Voth (2007)]. These glass-cutters, instrument makers, and specialists in fine mechanics were crucial in turning ideas into working prototypes, or existing machines into reliable capital equipment.

Industrialization occurs in our model in the following way: Incomes fluctuate around their long-run trend, pinned down by the demographic regime in the pre-industrial era. Technology improves but slowly through the use of capital itself – the more manufacturing activity there is, the more scope there is for improving and refining existing designs. The higher pre-industrial incomes, the greater the chance that a positive, persistent shock leads to a large increase in manufacturing output. The higher manufacturing output, the more capital-intensive production overall becomes – and the greater the scope for an acceleration of productivity growth because of growing differentiation of capital inputs. This setup resolves the apparent incompatibility of endogenous growth models with the history of technology which was emphasized by Crafts [1995]. One of the key criticisms of long-run growth models by economic historians has been that they often imply important and large scale effects – and that countries with bigger markets should have industrialized first [Crafts (1995)]. Yet the richest countries in early modern Europe were typically small, as was Britain for most of the period before 1750. We deliberately avoid these pitfalls by offering a mechanism for industrialization that does not presume that bigger countries have an automatic advantage.[2]

[2] Unified growth theory models in the spirit of Galor-Weil do not predict that bigger countries should industrialize first. Rather,

Population grows in response to higher wages; positive shocks to income add to demographic pressures, but also increase the scope for the capital externality to work. Crucially, because of fertility limitation, Europe's birth rates never outpace the rate of capital accumulation. We argue that England in particular (and Europe in general) had a higher chance to undergo a transition because of the high initial starting incomes and a favorable demographic regime.

For our argument to hold, England had to be ahead of the rest of Europe – and Europe markedly ahead of the rest of the world – in terms of per capita income. This is not accepted with unanimity. Pomeranz (2000) argues that, in the Yangtze region in China, living standards were broadly similar with the most advanced regions in Europe, and that the "great divergence" between Asia and Europe was a result of industrialization. Broadberry and Gupta (2005a) have recently shown that Pomeranz's claims, even for the Yangtze area, are probably exaggerated. Allen (2005a) finds that because of low rice and grain prices, the standard of living in Asia and Europe was broadly similar. However, money wages were markedly lower, and the relative price of manufacturing goods much higher. This is compatible with our interpretation, since it hinges on the purchasing power of income not dedicated to food.

1.3 The Model

This section sets out the basic setup of our model. The economy is composed of infinitely-lived, identical households whose members work, consume, invest, and procreate. Households choose consumption and saving to maximize their dynastic utility, subject to an intertemporal budget constraint. We consider a representative household of size N. In the following, we will refer to N as population and to household members as consumers or individuals. Current family members expect N to grow at the rate $\gamma_N(\cdot)$ because of the net influences of fertility and mortality, depending on consumption. In every period the economy produces two types of consumption goods: food and manufacturing products – and investment goods in the form of capital varieties. Output is produced using land, labor, and the accumulated stock of capital varieties. Consumers' preferences are non-homothetic: Representing Engel's law, the share of manufacturing expenditures grows with income. Below, we describe each of these elements of our model in turn.

1.3.1 Consumers

Each household member supplies one unit of labor in every period. Families use their income for investment, and to consume an agricultural good (c_A) and a manufactured good (c_M). Households maximize their expected life-time utility in a two-stage decision. In an intertemporal optimization problem, they decide upon consumption expenditure per household member in a given period t, e_t. In the second stage, the intra-temporal optimization, each individual takes e_t as given and maximizes instantaneous utility. We consider the second stage first. The corresponding budget constraint is $c_{A,t} + p_{M,t}c_{M,t} \leq e_t$, where $p_{M,t}$ is the price of a manufactured good. The agricultural good serves as the numeraire. Before they begin to demand manufactured goods, individuals need to consume a minimum quantity of food, $\underline{c}$. Preferences

their unit of observation is the world, and they simply assume that the partial derivative of technogical change with respect to population size is positive. At this level of aggregation, economic historians cannot disagree. The difficulty appears to be that for a model that captures cross-sectional differences, factors other than size must be important, and it is these factors that we try to capture.

take the Stone-Geary form and imply the composite consumption index:

$$u(c_A, c_M) = (c_A - \underline{c})^{\alpha} c_M^{1-\alpha} \tag{1.1}$$

Given e_t, consumers maximize (2.1) subject to the budget constraint. This yields the following equation for the expenditure share on agricultural products:

$$\frac{c_{A,t}}{e_t} = \alpha + (1-\alpha)\left(\frac{\underline{c}}{e_t}\right) \tag{1.2}$$

In a poor economy, where income is just enough to ensure subsistence consumption $\underline{c}$, all expenditure goes to food. As people become richer and e_t grows, the share of spending on food falls, in line with Engel's law. For very high levels of expenditure, $\underline{c}/e_t$ converges to zero and the agricultural expenditure share converges to α, which can thus be considered the food expenditure share in a rich economy.

We now turn to intertemporal optimization. First, we derive the indirect utility of consumers from (2.1) and (1.2):

$$\widetilde{u}(p_{M,t}, e_t) = \left(\frac{1}{p_{M,t}}\right)^{1-\alpha} \alpha^{\alpha} (1-\alpha)^{1-\alpha} (e_t - \underline{c}) \tag{1.3}$$

We use this result to set up the intertemporal optimization problem. The representative household maximizes expected dynastic utility subject to the intertemporal budget constraint:

$$\max_{\{k_{t+1}\}_{t=0}^{\infty}} E_0 \sum_{t=0}^{\infty} \beta^t \frac{[\widetilde{u}(p_{M,t}, e_t)]^{1-\psi} - 1}{1-\psi} N_t \tag{1.4}$$

$$\text{s.t.} \quad (1+\gamma_{N,t})k_{t+1} = (1-\delta)k_t + \tfrac{1}{p_{K,t}}(y_t - e_t)$$
$$y_t = w_t + r_{L,t} l + R_{K,t} p_{K,t} k_t$$

where y_t, e_t, k_t, and l are per-capita income, consumption expenditure, capital, and land, respectively, and $p_{K,t}$ is the price of capital.[3] Factor returns are the gross capital interest rate R_K, wage w, and the land rental rate r_L. Capital depreciates at rate δ; $\beta \in (0,1)$ is the households' discount rate, and $1/\psi$ gives the (constant) elasticity of intertemporal substitution, where $\psi \geq 1$. Since families take care of the welfare and resources of their prospective descendants, individual instantaneous utility $\left(\widetilde{u}(\cdot)^{1-\psi} - 1\right)/(1-\psi)$ is multiplied by household members N. Households take population growth $\gamma_{N,t}$ as given when optimizing. Together with the budget constraint, (1.4) yields the Euler equation

$$\left(\frac{1}{e_t - \underline{c}}\right)^{\psi} = \beta E_t \left[\left(\frac{p_{K,t+1}}{p_{K,t}}\right)\left(\frac{p_{M,t+1}}{p_{M,t}}\right)^{(1-\alpha)(\psi-1)}\left(\frac{1}{e_{t+1} - \underline{c}}\right)^{\psi}(1 + R_{K,t+1} - \delta)\right] \tag{1.5}$$

The Euler equation is the standard one, except for the two price terms on the right-hand side. If the price of manufactured goods increases, consumption in the next period will be more expensive. If the elasticity of intertemporal substitution is low (i.e., $\psi > 1$), the income effect will outweigh the substitution effect, and consumption e_t will be lower. If the price of capital p_K is expected to increase, investment is shifted from the future to the present, also lowering today's consumption. We use policy function iteration to solve the Euler equation, as described in Appendix A.6.

[3] In our model capital is the collection of varieties (machines). Thus, total capital $K = kN$ is equal to the integral over all capital varieties used in the economy. We provide a formal description of the capital stock below.

1.3.2 Production

Firms produce both capital and final goods. The latter are either agricultural or manufactured, are homogenous, and are produced under perfect competition. Capital is non-homogenous. It comes in many varieties that are produced monopolistically subject to increasing returns. The efficiency of production depends on the number of capital goods varieties. Free entry in the capital-goods producing sector ensures that, in equilibrium, there are no profits.

1.3.3 Final goods

Final sector firms use labor N, land L, and capital in the form of varieties $j \in [0, J]$ to produce their output. The agricultural production function is

$$Y_A = A_A \left[\int_0^J \nu_A(j)^{\frac{1}{1+\epsilon}} dj \right]^{\phi(1+\epsilon)} N_A^{\mu} L^{1-\phi-\mu} \tag{1.6}$$

where A_A is a productivity parameter in agriculture and $\nu_A(j)$ is the amount of capital variety j used for agricultural production in a representative final sector firm. Productivity fluctuates over time: $A_{A,t} = z_t \underline{A}_{A,t}$, where the component z_t represents a shock with mean one. The shock parameter z_t follows the AR(1) process $\ln z_t = \theta \ln z_{t-1} + \varepsilon_t$ with autocorrelation θ and $\varepsilon_t \sim N\left(0, \sigma_\varepsilon^2\right)$. To capture the growth of agricultural productivity over the long term even before the Industrial Revolution, we let the efficiency parameter grow at rate γ_A, such that $\underline{A}_{A,t+1} = (1+\gamma_A)\, \underline{A}_{A,t}$. The shocks ε_t should be interpreted as caused by weather conditions rather than changes in technology [as in Gilboy 1932].[4] Production becomes more efficient if more varieties of capital goods j are available. These enter with the (constant) elasticity of substitution $(1+\epsilon)/\epsilon$, where $\epsilon > 0$. Land is a fixed factor of production.

Manufacturing production is given by

$$Y_M = A_M \left[\int_0^J \nu_M(j)^{\frac{1}{1+\epsilon}} dj \right]^{\eta(1+\epsilon)} N_M^{1-\eta} \tag{1.7}$$

where A_M is a productivity parameter, and $\nu_M(j)$ is the amount of capital variety j used to produce manufacturing output in a representative final sector firm.[5]

1.3.4 Capital

Technological progress takes the form of a growing variety of machines available for production. There are j types of capital. Each of them allows a firm to perform a specific task. The more specialized machinery is, the higher productivity in final goods production.[6] As the number of varieties grows, machines that are better-suited to each production task become available.

[4] The abundance or shortage of seed as well as the effect of storage on price in periods following good or bad harvests causes the autocorrelation of output. Cf. McCloskey and Nash (1984).

[5] Due to constant returns in final production, we can assume without loss of generality that final sector firms are identical and have mass one. Individual firms' output, Y_A and Y_M, and factor inputs are then equal to aggregate output and input in the final sector. Thus, a final sector firm represents aggregate final production.

[6] In the symmetric equilibrium, $\nu_M(j) = \bar{\nu}_M, \forall j$, and thus $Y_M = A_M J^{\eta\epsilon} (J\bar{\nu}_M)^{\eta} N_M^{1-\eta}$ Consequently, for a given amount of capital $J\bar{\nu}_M$, productivity is increasing in the number of available capital varieties J, and the extent of this externality is given by $\eta\epsilon$. Similarly, for agriculture production, the extent of the externality is given by $\phi\epsilon$.

Producers borrow capital from consumers, and pay interest at rate $r_K = R_K - \delta$. Producers replace depreciated capital while production occurs.[7] The price of a variety is $p(j)$. There are $\nu(j)$ items of type j machines available. Representative final sector firms then use $\nu_A(j)$ and $\nu_M(j)$ machines of type j for food and manufacturing production, respectively. We assume that the subset of varieties that break (depreciate) in a given period is the interval $[(1-\delta)J, J]$ of capital varieties.[8] The mass δJ of broken machines is replaced by producers while production occurs.

To start production, capital variety producers need to pay up-front cost F. Capital variety producer $\widetilde{\jmath}$ uses technology

$$\nu\left(\widetilde{\jmath}\right) = A_J \left[\int_0^J \nu_{\widetilde{\jmath}}(j)^{\frac{1}{1+\epsilon}} dj\right]^{\eta(1+\epsilon)} N_{\widetilde{\jmath}}^{1-\eta} \quad - \quad F \tag{1.8}$$

where A_J is a constant productivity parameter. Note that j refers to machines existing in a given period, whereas $\widetilde{\jmath}$ stands for capital varieties that are currently produced as investment goods, becoming available for production in the next period. Like final sector firms, capital producers profit from a wider range of available capital inputs.[9]

Because of fixed costs, each capital variety is produced by a single firm. Since capital varieties are imperfect substitutes, their producers have monopolistic power. However, free market entry implies that each producer just recovers his unit cost and makes zero profits. We show in Appendix A.1 that in equilibrium each firm produces the same, fixed amount of capital varieties, given by F/ϵ. Increasing investment leads to an extension of the range of capital varieties, while leaving the amount $\nu(j)$ of each variety unchanged. This, together with symmetry in equilibrium, allows us to derive a simplified aggregate externality representation of the model, where investment goods are produced in the manufacturing sector.

1.3.5 Model Representation with Aggregate Externalities

We show in Appendix A.4 that the production side of our model can be simplified to a two-sector model with externalities of aggregate capital in the style of Romer (1990). Technology is then given by

$$Y_A = A_A K^{\phi\epsilon} K_A^{\phi} N_A^{\mu} L^{1-\phi-\mu} \tag{1.9}$$

$$Y_M = A_M K^{\eta\epsilon} K_M^{\eta} N_M^{1-\eta} \tag{1.10}$$

where we introduced a more convenient notation for capital: $K_A \equiv J\nu_A$ and $K_M \equiv J\nu_M$, representing the capital used by a representative firm in the respective sector. Investment, i.e., new capital varieties, are produced by the manufacturing sector, and the price of capital, p_K, is equal to the price of manufacturing output, p_M.[10] The productivity-enhancing effect of an increased variety of capital inputs is obvious in

[7] This assumption becomes important in our simulation. It avoids that the capital stock falls until it finally reaches zero when consumers live at the subsistence level for a long time (Malthusian trap).

[8] A simple way to motivate this assumption is to think of machines $j \in [0, J]$ as being ordered by age, with higher-j subsets representing older machines. Due to their long use, or because of being incompatible with new machines, the highest-j subset with mass δ breaks or becomes useless in each period, and is immediately repaired or replaced by machines of equal quality. New machines fill up the interval from below, increasing J, but leaving the age-ordering unchanged.

[9] We deliberately deviate from the standard setup to simplify our analysis below, where we derive the model representation with two sectors and an aggregate externality.

[10] Since we have $p_K = p_M$ in the simplified model, the price terms in the Euler equation (1.5) simplify to $\left(p_{K,t+1}/p_{K,t}\right)^{\psi(1-\alpha)+\alpha}$, where the exponent is always positive.

these standard Cobb-Douglas production functions. For a given $J = K$, the aggregate externality is the larger the larger the capital share (ϕ or η) and the larger ϵ (i.e., the smaller the elasticity of substitution among capital varieties).

1.3.6 Equilibrium and Industrialization

Equilibrium in our model is a sequence of factor prices, goods prices, and quantities that satisfies the intertemporal and intra-temporal optimization problems for consumers and firms.[11] To fix ideas and show how industrialization happens in our model, we first present a simulation without consumption-dependent population dynamics. That is, we run our model with a positive constant birth rate and without shocks, such that all individuals survive. The next section explores how population dynamics – based on consumption-dependent fertility decisions and positive Malthusian checks in crisis periods – modify our results.

1.3.7 Equilibrium and Industrialization without Population Dynamics

In this section we keep population growth constant in order to isolate the role of consumption preferences (structural change) and aggregate capital externality. We show that even with this reduced-form model we are able to replicate two important stylized facts of the Industrial Revolution in England – the initially small, but accelerating growth of industry output and structural change, i.e., an increasing share of industry in GDP. We simulate the model with a constant birth rate, equal to the average rate in England 1700-1850, $\gamma_b = 0.8\%$. In a non-stochastic setup, these parameters imply that consumption never falls below subsistence such that all individuals survive. We thus have neither a preventive (via birth rates) nor a positive (via death rates) Malthusian check. The corresponding results are shown in figure 1.1.

Our simulation for England starts with the historical labor shares in agriculture and manufacturing in 1700 (77% and 23%, respectively).[12] Initially about half of manufacturing output is produced to replace depreciated capital, with the other half being used for consumption. Consumption exceeds the subsistence level so that all individuals survive and net population growth equals the birth rate ($\gamma_N = \gamma_b$). Figure 1.1 shows that our model, even with constant birth rates, replicates the low, increasing growth rates observed in 18^{th} century England. Growth is driven by the exogenous productivity progress in agriculture and by endogenous capital accumulation. Technological progress is fast enough to compensate the constant population growth of 0.8%, so that p.c. income increases.[13] Per capita consumption of agriculture grows much slower than p.c. output of manufacturing. This is explained by two mechanisms: First, as p.c. income grows, consumption expenditure shares shift from agriculture to manufacturing (as shown in the lower right panel). Once this transition is completed, industry growth rates fall but remain above those of agriculture, which is explained by the second mechanism: due to their larger capital share, manufacturing firms profit relatively more from the aggregate externality. This is reflected in the upper right panel: Initially, agricultural and manufacturing TFP grow in tandem – the larger growth rate of p.c. industry

[11] A formal definition of the equilibrium is given in Appendix A.3.

[12] We use the same parameter values as in the full, calibrated model. Our conclusions with regard to structural change and the role of capital externalities are robust with respect to the choice of parameters.

[13] This would not be the case if birth rates were substantially larger, since then p.c. capital would diminish at a rate that even the aggregate externality would not be able to compensate.

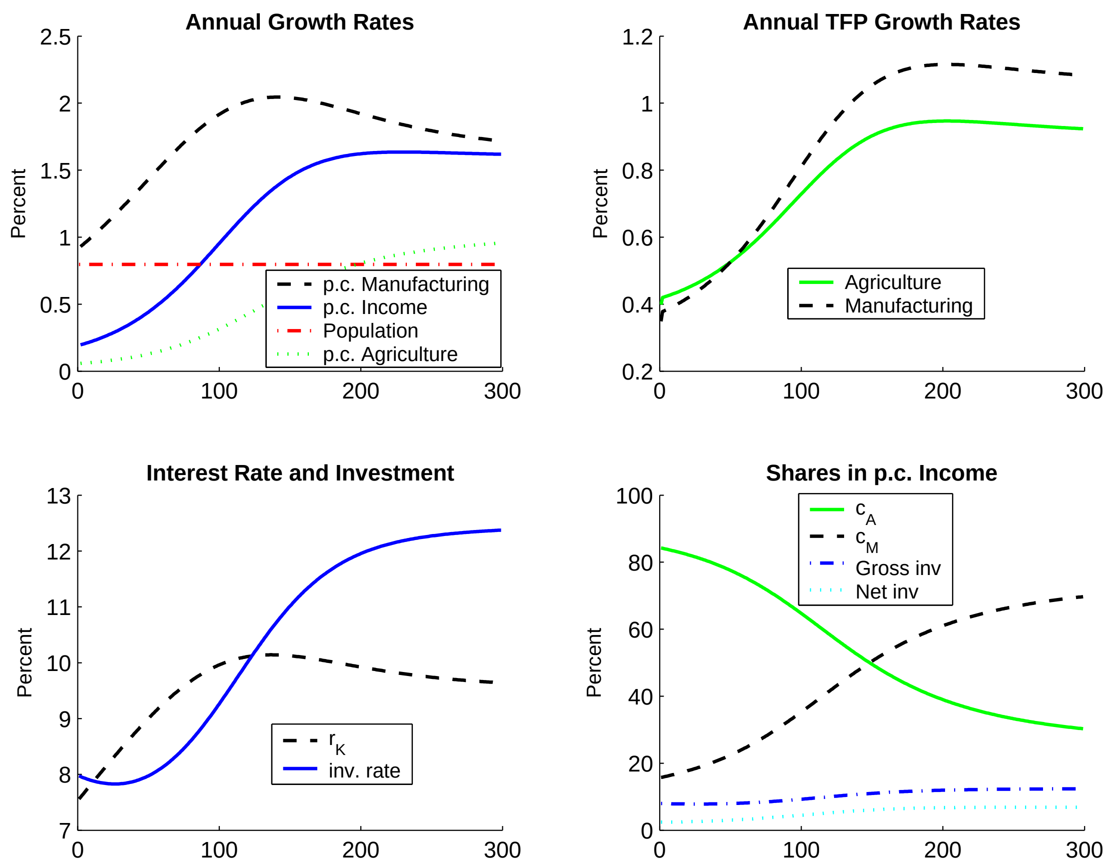


Figure 1.1: Simulation Results with Constant Population Growth

output is thus initially solely due to its increasing demand share. When structural change comes to a standstill, TFP and all other growth rates stabilize at constant levels, with manufacturing TFP augmenting faster than agricultural TFP. The investment rate is low initially because p.c. consumption is close to subsistence. Investment then responds positively to growing income and interest rates. Eventually, when p.c. consumption has grown to a level well beyond subsistence and interest rates level off, investment rates stabilize at a higher level.

1.3.8 Open Economy Considerations

So far, we have assumed that the UK was a closed economy, with domestic conditions driving industrialization. Because of its role as a trading nation, this needs to be justified in the British case. Before it started to manufacture cotton goods with new technology, for example, Britain imported many of them from India.[14] Eventually, Britain exported cotton goods and the like on a grand scale. Traditional interpretations of the importance of demand have assigned an important role to exports [Cole (1973) Gilboy (1932)]. This could also affect the logic of our argument – in some open economy models, lower initial

[14] In the 1750s, Indian cotton piece exports to Britain were five times higher than British exports. Exports from India to Britain only collapsed after 1810 [Broadberry and Gupta (2005b), table 6].

agricultural productivity can increase the probability of industrialization since wages (and thus, prices of exports) are lower [Matsuyama (1991)]. Here, we discuss how adding foreign demand and supply would change our basic setup.

The fact that British industrialization in cotton textiles replaced exports from India as such does not fundamentally alter our conclusions. Rising manufacturing productivity has two consequences in the model: higher p.c. income and lower prices. Both increase the demand for manufacturing output (the former through Engel's Law). In an open-economy framework, the price effect is even larger, because imports are replaced by home production and/or due to growing international demand. The falling relative price of manufacturing would also be expected to result in growing food imports.

An open economy setup, especially for the 18^{th} century, must take into account the high cost of transportation. These made it (i) easier to replace the Indian competition in the UK and (ii) isolated the Indian producers from UK competition for some of the time.[15] Table 1.2 shows that between 1750 and 1851, the share of exports - mainly of manufacturing products - in national output grew from about 15% to 20%. As Mokyr (1977) stressed, there is no evidence that exports grew sufficiently rapidly to kick-start industrialization. We conclude that our closed-economy model can serve as a reasonable approximation.[16]

Table 1.2: International Trade in England 1700-1851

Year	Exports / Output	Manufactures Exports / Output	Food Imports / Output
1750	14.6*	11.0	4.5
1801	15.7	13.8	6.1
1831	14.3	13.0	3.9
1851	19.6	15.9	7.2

Source: Crafts (1985), Table 6.6 and 7.1.Authors'calculation assuming balanced trade. All numbers in per cent.
* Number for 1760.

1.3.9 Inequality

To capture one particular feature of the pre-modern world highlighted by Fogel (1994), we also consider the economic contribution of the bottom 20% of the income distribution. According to Fogel, in eighteenth-century France, the poorest 20% did not receive enough food to perform more than a few hours of work a day. We model such an outcome by assuming that, if average consumption falls below subsistence, members of the workforce that will die because of malnutrition will also not be able to work.

[15] Initially, Indian exports become uncompetitive in Britain as the UK switches to industrialized production. Home production in India remains competitive while transport costs raise the price of UK cotton goods there. Eventually, Indian production of cotton goods for home demand falls as UK imports become cheaper due to falling transport cost [Broadberry and Gupta (2005b)].

[16] Total output, Y, approximately quadrupled between 1750 (t=0) and 1850 (t=T) [Crafts (1985)]. From simple growth accounting, we have: $\frac{Y^T - Y^0}{Y^0} = \left[s_E^T \frac{Y^T}{Y^0} - s_E^0\right] + \left[(1 - s_E^T)\frac{Y^T}{Y^0} - (1 - s_E^0)\right]$ where the parentheses indicate output growth due to exports and domestic demand, respectively. Let the share of exports grow from $s_E^0 = 15\%$ to $s_E^T = 20\%$, as in table 1.2. Then, 78% of growth is due to domestic demand, while exports account only for 22%.

This is clearly too optimistic – even without starvation, many members of the workforce will be malnourished. When harvest failures occur, the effective workforce will shrink – except in England, which provided generous support to the poor via outdoor relief, especially during the years of high prices in the late eighteenth century. In the other two countries we consider – France and China – we assume that there is no redistribution.

1.3.10 Population Dynamics

Having summarized the basic properties of the economy, we now add population dynamics. At low levels of productivity, the economy is Malthusian. As agricultural productivity increases, population expands. As land-labor ratios fall, living standards decline and return to their earlier level. If times are bad, starvation can cause sharp declines in population size. We show how certain features of the demographic regime can make the escape from the Malthusian trap possible. In particular, we demonstrate how a low-pressure regime with limited fertility increases the chances for sustained growth.

The size of the representative household (or population) N increases by a factor of $g_b(\cdot)$ at the end of each period:

$$N^*_{t+1} = g_b N_t \tag{1.11}$$

where N^*_t is the beginning-of-period population, whereas N_t stands for the population that survived period t. The exact growth factor depends on the demographic regime. At one extreme ("high pressure regime"), we assume a constant birth rate $\overline{g}_b$. Here, population returns to equilibrium after negative shocks through more deaths (e.g., Malthus' positive check). Alternatively ("low pressure regime"), the birth rate depends positively on real consumption, $g_b(c_t)$.[17] This is because the European marriage pattern regulated population-wide fertility by changing marriage rates. In bad times, people married later, and fewer women ever married. Within marriage, there were no signs of fertility-limitation. In this way, population is balanced by the operation of both the positive and the preventive check.

We assume that if consumption per head falls below $\underline{c}$, only a subset of the population survives. The probability of survival depends on the severity of the nutrition crisis, measured by the ratio of c_t to $\underline{c}$:

$$g_s(c_t) = \frac{N_t}{N^*_t} = \min\left\{\frac{c_t}{\underline{c}}\ ,\ 1\right\} \tag{1.12}$$

With severe harvest failures, population falls, and starving individuals consume their capital. They die when they have exhausted it.[18]

It could be argued that population growth should only depend on income in terms of agricultural goods (as in Strulik 2006). We consider our approach more intuitive, since goods produced in urban centres were clearly an important part of the consumption bundle even for poor people (King 1997) before the Industrial Revolution, as reflected by urbanization rates. However, the basic mechanism enabling sustained growth is robust to changing the population growth function in the manner of Strulik (2006). Since fertility responds only to one part of income, population growth is slower. The positive externality

[17] Concretely, c_t denotes per-capita consumption of agriculture and manufacturing goods, that is, $c_t = c_{A,t} + c_{M,t}$.

[18] Diamond (2004) describes how the Norse colony in Greenland collapsed after years of worsening climatic conditions, until farmers started to eat their calves and seed corn.

has a smaller effect. Hence, TFP and output growth also slow down. However, industrializations still occur with a high frequency.[19]

Population growth $\gamma_{N,t}$ is a function of economic conditions:

$$\gamma_{N,t} = \frac{N^*_{t+1} - N^*_t}{N^*_t} = g_b g_s(c_t) - 1 \tag{1.13}$$

where g_b depends on c_t or is a positive constant.

The birth function $g_b(\cdot)$ is crucial for the escape from the Malthusian trap.[20] If birth rates at low levels of consumption are also low, and the response of births to improving conditions is small, productivity growth can translate into growth of per-capita income (despite the fact that population grows). This will be the case if $g_b(\cdot)$ is relatively flat at $\underline{c}$.

Where $g_b(\cdot)$ is a positive constant, escaping the Malthusian trap is nearly impossible. If the constant birth rate g_b exceeds productivity growth, resources are not sufficient to nurture everyone and the surviving population remains trapped at the subsistence level.[21] We will from now on use the full model, with population dynamics. Next, we describe the economic effects of demographic interactions, contrasting the "low pressure" and the "high pressure" regimes. In this setup, we show how fertility limitation helps the escape from the Malthusian trap.

Figure 1.2 shows population growth as a function of capital per head (k) – in the left panel for the low-pressure regime and in the right panel for the high-pressure regime. Capital stock per head corresponds to a certain level of per capita income, given a certain level of TFP. As incomes and consumption improve, birth rates γ_b increase in the low-pressure regime, while they are constant in the high-pressure regime. Above point A, income rises with k such that death rates (given by $\gamma_b - \gamma_N$) dwindle to zero. The solid black line shows the gross rate of capital formation, inv/k, where real investment is $inv = (y - e)/(p_K)$.[22] The growth rate of capital stock per capita is given by the difference between inv/k and effective depreciation ($\delta + \gamma_N$). In equilibrium with constant k, the capital-diluting effects of population growth and depreciation offset each other: $(\delta + \gamma_N)k = (y - e)/p_K$.

We begin by analyzing the low-pressure regime. To the left of point A, consumption is below subsistence ($c < \underline{c}$), and due to the crisis no new individuals are born ($\gamma_b = 0$). Investment just replaces depreciation.[23] Net population growth γ_N is negative such that the increasing land-labor ratio implied by falling population finally drives the economy back to an equilibrium at point A. At point A, consumption is at subsistence ($c = \underline{c}$); the birth rate is zero. Point A is an unstable equilibrium. For higher levels of k, incomes improve and investment rises. Eventually, declining marginal returns to capital force down the 'inv/k' curve. The new (stable) equilibrium is point B, which combines constant k and a growing population.

In the high pressure regime, the economy behaves differently. The right panel of figure 1.2 depicts the interactions of demographic growth, investment, and output. For low levels of capital, there is also

[19]The growth rate of output per capita over 150 periods with the Strulik assumptions is 0.34% instead of 0.56% in the deterministic baseline simulation.

[20]See Appendix A.7 for our calibration of the birth function for England.

[21]In our calibrated model for China in 1700, for example, the constant birth rate is 4%, while deaths occur with rate 3.2%, implying a net population growth of 0.8% p.a.

[22]This is gross of depreciation.

[23]This follows from our assumption that producers immediately replace depreciated capital varieties. Without this assumption consumers would choose not to repair the capital stock and even consume out of it if consumption falls below subsistence.

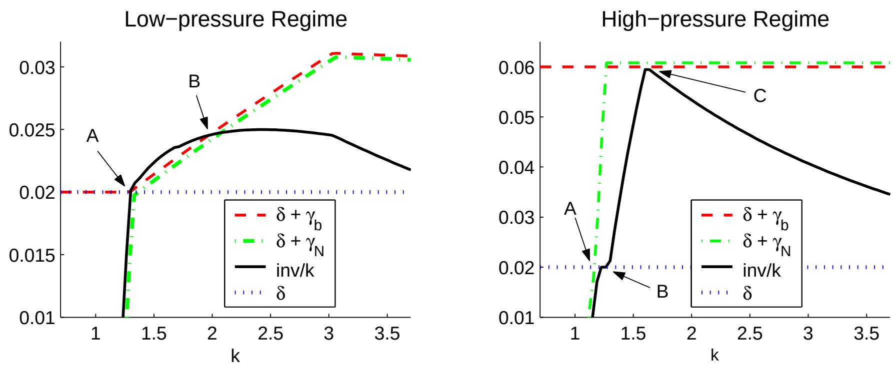


Figure 1.2: Population Dynamics for England and China

starvation, as in England. Point A now is a stable equilibrium with $c < \underline{c}$, and birth rates that are offset by death rates. However, with capital slightly higher than at point A, death rates fall quickly until the economy reaches point B, where $c = \underline{c}$.[24] Now, death rates are zero, and demographic growth becomes very fast. Consumers respond to this rapid population growth by investing massively, in order to ensure minimal consumption tomorrow, when they expect to share their income with many others. This explains the steep slope of the 'inv/k' curve to the right of point B. However, despite saving all income above subsistence, demographic growth is too rapid – capital-labor ratios fall, driving the economy back to point A. If the economy reaches point C, capital-labor ratios stabilize, as the capital stock expands at the same rate as population. However, point C is not a stable equilibrium since a small negative shock will drive the economy back to point A. To the right of C, investment falls rapidly, as marginal returns decline and saving rates reduce.

Only the low pressure regime is likely to generate endogenous TFP growth. At point B in the low-pressure regime, total capital is growing with population. Because of the aggregate externality, this generates TFP growth. In figure 1.2 this would be equivalent to a shift up and to the left of the 'inv/k' line – for any given capital stock, incomes are now higher. There is also an outward shift of the birth schedule, since higher incomes stimulate higher birth rates and sustain a larger population at the same p.c. capital level. The combined effect under our calibration leads to a point B' that is markedly higher, and further to the right – TFP growth produces a new equilibrium B' that is more capital intensive, has higher incomes, and more rapid population growth. This explains the gradual acceleration of growth rates in the low pressure regime.

In the high pressure regime, endogenous growth is not impossible but highly unlikely. Higher TFP simply shifts the investment schedule to the left – for any given level of capital, potential consumption is higher, but so is population growth. Higher productivity leads to a bigger population, with unchanged income at A. If the (constant) birth rate under the high-pressure regime was low enough, growth could occur, because the investment schedule would eventually cross the line given by $\delta + \gamma_N$. This would create a stable equilibrium point C, similar to point B in the low-pressure regime. The maximum rate of population increase that does not exhaust investment possibilities varies with starting conditions. In

[24]Between A and B, net investment is zero because consumption is below subsistence.

our calibration, a country with an initial non-agricultural labor share of 23% (equivalent to Britain's in 1700) could have sustained population growth rates of up to 3.7 percent because of high initial income; a country with only 10% in non-agricultural occupations (as China in 1700) could not have coped with rates higher than 0.6 percent without foregoing its chances to industrialize.[25]

1.4 Calibration and Simulation Results

In this section we explain the calibration of our model, and simulate it with and without shocks to agricultural productivity. We then derive the probability of industrialization in England, France, and China. In addition, we illustrate what would have happened to the English economy had it operated under a high-pressure demographic regime instead. Finally, we simulate the model without the kind of redistribution that the Poor Law provided.

1.4.1 Calibration

We normalize initial population of England to unity ($N_0 = 1$) and choose land $L = 8$ such that its rental rate is 5%. We choose initial agricultural TFP and aggregate capital to match the historical labor share in agriculture of 77% in 1700.[26] Aggregate capital K influences TFP via the externality. In order to identify the initial conditions for $A_{A,0}$ and K_0, we re-normalize the production functions, dividing by $K_0^{\phi\epsilon}$ in agriculture and by $K_0^{\eta\epsilon}$ in manufacturing. This means that the aggregate externality term has value one in the initial period.[27] We choose A_M such that the price of manufacturing products is double the price of agriculture products, i.e., $p_M = 2$.[28] This procedure gives $A_{A,0} = 0.517$ and $A_M = 0.359$. Given these parameters, we derive a low level of capital, K_{min}, at which consumption is at the subsistence level ($c = \underline{c}$). Below this level, only agricultural goods are consumed, and aggregate capital does not influence TFP. The externality works only if $K \geq K_{min}$.[29] In other words, it is not before the "wave of gadgets" [Ashton (1949)] arrives that the aggregate externality begins to matter quantitatively.

In the centuries before 1700, labor productivity grew at an average rate of 0.15% per year (Galor 2005). Because agriculture was the dominant sector, we assume an exogenous growth rate of TFP growth in the sector of $\gamma_A = 0.15\%$.

The magnitude and persistence of shocks in the agricultural sector is derived from real wage data for England, 1600-1780 [Wrigley and Schofield (1997)]. With fixed labor supply and agriculture the dominant sector, these productivity shocks have an immediate knock-on effect on real wages in the economy. This is especially true since wages were largely fixed in nominal terms, and most of the variation in the Phelps-Brown/Hopkins wage series results from changes in agricultural prices [Wrigley and Schofield

[25] These are the results for non-stochastic simulations. In calibrations with shocks, there would be a distribution of industrialization outcomes for each demographic growth rate.

[26] We derive this figure from Craft's (1985) original numbers by leaving out other sectors than agriculture and manufacturing and re-normalizing the sum of these two sectors' labor shares to unity.

[27] This normalization does not change any of the features of our model. In fact, dividing K by K_0 is equivalent to re-defining A in the production function. For example, let the original production function be $Y_M = A_M^* K^{\eta\epsilon} K_M^{\eta} N_M^{1-\eta}$. Then choose A_M such that $A_M^* = A_M / K_0^{\eta\epsilon}$. This gives the new production function $Y_M = A_M (K/K_0)^{\eta\epsilon} K_M^{\eta} N_M^{1-\eta}$.

[28] Different values of this parameter change our results only slightly. They do so at all because $p_M = p_K$, and a different price of capital implies a different real capital stock.

[29] The aggregate externality thus takes on the values $[\max\{\frac{K}{K_0}, \frac{K_{min}}{K_0}\}]^{\phi\epsilon}$, in agriculture and $[\max\{\frac{K}{K_0}, \frac{K_{min}}{K_0}\}]^{\eta\epsilon}$ in manufacturing.

(1997)]. We therefore use the wage z_t as an indicator of the size of shocks. Figure 1.3 shows the real wage index and the corresponding Hodrick-Prescott–trend.[30]

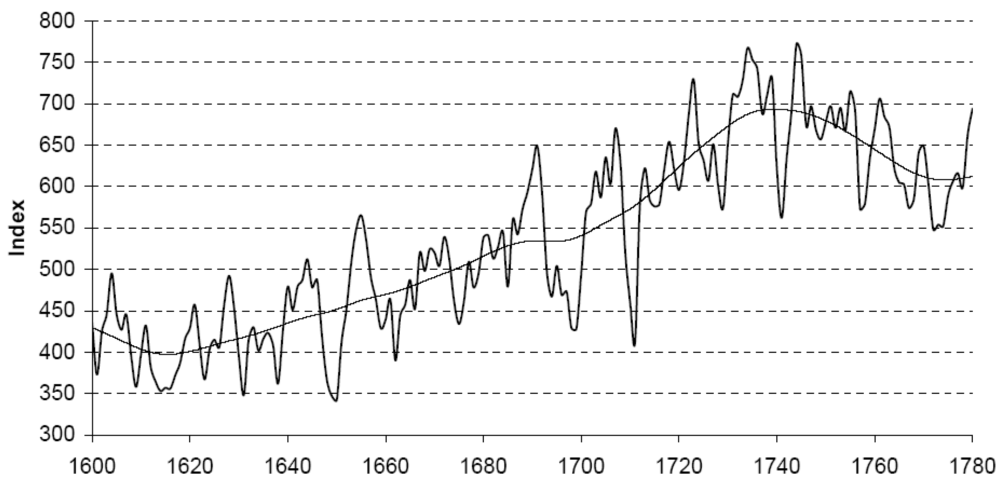


Figure 1.3: Real Wage Fluctuation and Trend

The magnitude of shocks is derived from analyzing the autocorrelation of wages. We estimate $\ln z_t = \theta \ln z_{t-1} + \varepsilon_t$, which produces $\theta = 0.60$ (t=10.15) and $\sigma_\varepsilon = 0.075$. The autocorrelation of shocks is high, and the series is volatile.

For the baseline model, we calibrate the parameters (μ, ϕ, η, ϵ) to fit average factor shares for the period 1700-1850. In agriculture, we use $\mu = 0.4$ for labor, $\phi = 0.25$ for capital, and the remaining 0.35 for land. This is similar to the 40-20-40 split suggested by Crafts (1985), and is almost identical with the average in Stokey's (2001) two calibrations. In manufacturing, we use a capital share of $\eta = 0.35$.[31]

We normalize the minimum food consumption $\underline{c}$ to unity. For low income levels, equation (1.2) implies that all expenditure goes to agriculture. With higher incomes, the expenditure share converges to α. We take expenditure data from Crafts (1985), using the same re-normalization as for labor shares. The agriculture consumption share falls from 65% in the 18^{th} to 30% at the end of the 19^{th} century. We thus use $\alpha = 0.3$. Next, we need to choose ψ, i.e., the inverse of the intertemporal elasticity of substitution. In the literature, values between 1 and 4 have been used. We employ $\psi = 1$, which implies log-utility, because this matches the elastic supply of savings during the Industrial Revolution.[32] In order to capture the low initial share of investment (4% in 1700, 6% in 1760, taken from Crafts 1985, table 4.1), we need a low discount factor, and use $\beta = 0.93$ and depreciation rate $\delta = 0.02$.

The aggregate externality plays a central role in our model. The extent of the externality is given by $\phi\epsilon$ in agriculture and by $\eta\epsilon$ in manufacturing production. In manufacturing, total factor productivity is given by $A_M(K/K_0)^{\eta\epsilon}$, where the first term and K_0 are constant. Growth of manufacturing TFP is

$$\gamma_{T,M} = \eta\epsilon\, \gamma_K \tag{1.14}$$

[30] The standard deviation of real wages is very similar to the standard deviation of agricultural output in later years.

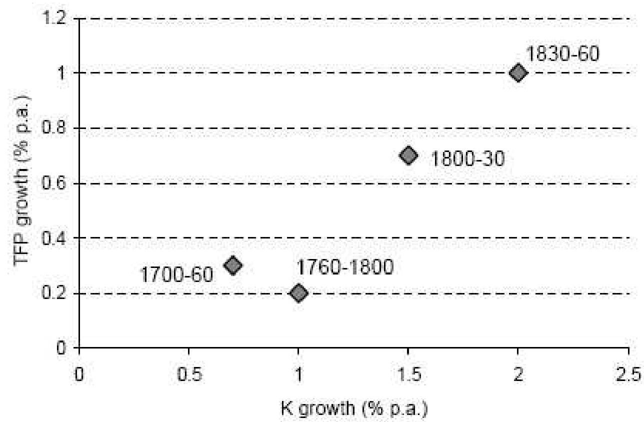
[31] Stokey (2001) uses a calibration for an energy-capital aggregate with the average share of 0.4.

[32] The higher intertemporal elasticity of substitution implied by the smaller ψ means that consumers' savings react more elastically to changes in the interest rate. On the high elasticity of savings, see Allen (2005b).

Total factor productivity in agriculture is determined by $A_A(K/K_0)^{\phi\epsilon}$, where the first term grows at the exogenous rate γ_A:

$$\gamma_{T,A} = \gamma_A + \phi\epsilon\,\gamma_K \tag{1.15}$$

Crafts (1985) provides growth accounting figures for England, 1700-1860. We present the corresponding TFP and aggregate capital growth rates in Figure 3. If the aggregate externality link from capital to TFP in our model represents historical facts, we would expect a linear relationship between the growth rates of the two variables. Figure 1.4 lends some support to this supposition.[33]



Source: Derived from Crafts (1985) and Crafts and Harley (1992).

Figure 1.4: Annual Growth Rates of TFP and Aggregate Capital

Average annual growth rates are $\overline{\gamma}_K = 1.17\%$ and $\overline{\gamma}_T = 0.48\%$ for capital and TFP, respectively. There is no agreement in the literature as to whether productivity growth in agriculture was faster, slower or equal to productivity growth in modern sectors. For example, Crafts (1985: 70-89) concluded that productivity growth in agriculture was rapid, and in some periods surpassed manufacturing productivity growth. On the other hand, Clark (2003b) argued that they took a long time to materialize. We therefore assume that the growth of labor productivity was broadly speaking the same in manufacturing and agriculture. Thus, aggregate TFP growth is equal to sectoral TFP growth, and we can estimate the relationship (1.14) using the data represented in figure 3. A weighted least-square estimation (with the length of periods serving as weights) without constant yields the estimate $\widehat{\eta\epsilon} = 0.44$ (t=7.26).[34] With $\eta = 0.35$, this implies $\epsilon = 1.25$, corresponding to an elasticity of substitution across capital varieties of 1.8. There is an easy way to check the consistency of this calibration with other calibrated variables: we use the observed $\overline{\gamma}_K$ and $\overline{\gamma}_T$ together with the calibrated γ_A, ϕ, and ϵ to check (1.15). The result is 0.51% on the right-hand side, which corresponds well to $\overline{\gamma}_T = 0.48\%$.[35] For the observed growth of aggregate capital 1700-1860, our calibration thus implies very similar TFP growth rates in manufacturing and agriculture, where the latter also includes an exogenous term.

[33] Of course, we do not claim here that our model is the only explanation of the relationship observed in the growth accounting data. In fact, the causality could also go the other way around – from exogenous TFP growth to capital accumulation. However, what matters for our calibration is the linearity of the relationship, while we suppose the direction of causality to be from K to TFP, along the main line of our argument relating to an increasing number of available capital varieties.

[34] Another possibility is to take average values instead of running a regression. The result is very similar: $\overline{\gamma}_K/\overline{\gamma}_T$ = 0.41.

[35] Our choice of the capital shares ϕ and η is crucial for this result.

We employ a birth schedule $g_b(c)$ based on the historically-observed co-movement with wages (cf. Figure A.1).[36] It is derived from fitting the empirical data with a spline regression, as described in detail in Appendix A.7. For the demographic regime with positive Malthusian check, g_b is a constant equal to the net birth rate.

We summarize the calibration parameters in Table 2.1.

Table 1.3: Baseline Calibration

Symbol	Interpretation	Value
Parameters		
α	Agriculture expenditure share	0.3
β	Consumer discount rate	0.93
ψ	CRRA utility parameter	1
ϕ	Capital share in agriculture	0.25
μ	Labor share in agriculture	0.4
η	Capital share in manufacturing	0.35
ϵ	Parameter for capital variety substitutability	1.25
$\underline{c}$	Subsistence food consumption	1
L	Land	8
δ	Capital depreciation rate	0.02
γ_A	Growth of agriculture technology	0.0015
θ	Autocorrelation of shocks to agriculture	0.6
σ_ε	Standard Deviation of shocks to agriculture	0.075
A_M	Manufacturing technology parameter	0.359
Initial Conditions		
N_0	Initial population	1
$A_{A,0}$	Initial agriculture technology parameter	0.517
K_0	Initial aggregate capital	1.718
K_{min}	Capital at which $c = \underline{c}$	1.308

1.4.2 The Industrial Revolution in England

How well does the calibrated version of our model fit the historical data for England? We start in 1700 and run the simulation for 150 years. Figure 1.5 compares the non-stochastic simulation and historical facts. Over the period as a whole, population triples, while real per capita income doubles – mainly due to the increase of manufacturing output. Importantly, growth rates of output and TFP are initially low but increase over time. The model does well in capturing one of the key characteristics highlighted by economic historians in recent years – the slow rate of productivity and output growth [Crafts and Harley (1992)]. Also, output of agricultural products increases only slightly in our model, in line with the historical record [Allen (1992), table 8.7].

[36] We use the data from Wrigley and Schofield (1997).

The behavior of population and real manufacturing output is captured well by the model, even if we overestimate the growth of the latter somewhat. Initially, investment mainly replaces depreciated capital. Even with a low depreciation rate of $\delta = 0.02$, this implies an investment share of about 6%, which exceeds the historical estimate for 1700.[37] Our simulation replicates the rise of the investment rate during the following decades, but falls short of its full extent. One possible reason is changes in δ. Depreciation rates may have increased over time because machines became increasingly complex and technological obsolesence rendered useable equipment unprofitable. Real investment per capita grows by a factor of 3.5, which is accounted for by an increasing investment rate, growing income, and a falling relative price of capital (dropping by 25%). Population growth peaks around 1820, which coincides with the historical facts. TFP in agriculture and manufacturing is growing at similar rates. Agriculture benefits from exogenous growth ($\gamma_A = 0.15\%$); manufacturing from the greater externality resulting from its higher capital share. Payments to land become less imporant in total output, while capital and labor gain about 5% each. Stokey (2001) shows that labor and capital gained a larger share of the pie, and that land lost about 10 percentage points of aggregate income – yet the gains for capital in our model are somewhat smaller than the historical record suggests.

Employment shares in agriculture and manufacturing fit the data well, while the model overestimates the income share of agriculture.[38] One reason for this is hidden unemployment in agriculture – many workers in the English fields in 1700 may have added little to output. With the beginning of the Industrial Revolution around 1780, many of these laborers migrated to the cities. For these later years, the fit with our model is markedly better. Finally, TFP growth in our simulation fits the actual data well.[39]

1.4.3 Sensitivity Analysis

In the following we provide robustness checks of our model. We start from the baseline calibration and sequentially change key parameters [similar to Lagerlöf (2006)]. The results are summarized in table 1.4. Our baseline used an exogenous rate of agricultural TFP growth, $\gamma_A = 0.15\%$ p.a. In the first sensitivity check, we set this to zero. In order to fit the observed relationship between capital and TFP growth (figure 1.4), we consequently re-calibrate ϵ, obtaining a higher value.[40] Thus, some of the growth that was previously exogenous is now the result of a stronger aggregate externality. With all other parameters identical with the baseline calibration, the simulation yields slower growth, capital accumulation, and structural change when $\gamma_A = 0$. The difference with our baseline is however relatively minor.

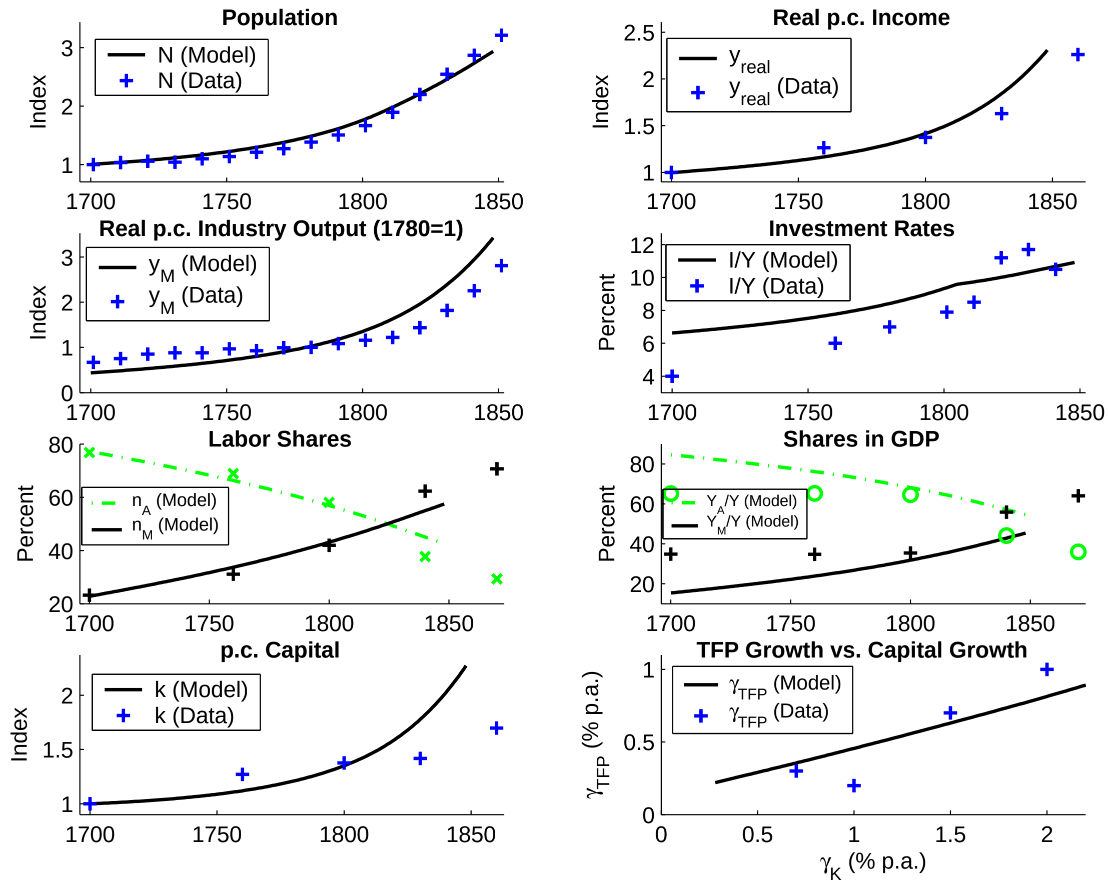
In the second alternative specification, we change the capital shares in agriculture ($\phi = 0.2$) and manufacturing ($\eta = 0.5$). This represents the ϕ suggested by Crafts (1985) and the η used by Stokey

[37] The corresponding equations are $I = \delta p_K K$ and $p_K R_K K = \tau Y$, where τ is the aggregate capital share. For $\tau \simeq 0.3$ and $R_K \simeq 0.1$ (the approximate values in 1700) this yields $I/Y = \delta\, \tau / R_K \simeq 0.06$.

[38] We derive the historical employment and income shares based on the numbers in Crafts (1985, p.62). We exclude the service sector, renormalize the percentages and interpolate to find the data for 1700 to 1860.

[39] The exception is the unusually low TFP in the late eighteenth century, when negative shocks such as the Napoleonic Wars may have made a big difference [Williamson (1984)], Temin and Voth (2005)].

[40] In the baseline calibration, equations (1.14) and (1.15) give $\gamma_{T,M} \simeq \gamma_{T,A} \simeq 0.51\%$ p.a. Thus, total TFP growth $\gamma_T \simeq 0.51$ in the baseline case. We now use this figure to derive ϵ for the case $\gamma_A = 0$. Given that the average share of agriculture in GDP was about 60% between 1700 and 1850 [abstracting from the service sector, which we do not model], the corresponding approximation $\gamma_T \simeq 0.6\gamma_{T,A} + 0.4\gamma_{T,M} = 0.6\phi\epsilon\, \overline{\gamma}_K + 0.4\eta\epsilon\, \overline{\gamma}_K$ implies $\epsilon \simeq 1.5$. Note that we cannot obtain $\gamma_{T,M} = \gamma_{T,A}$ if $\gamma_A = 0$ and $\phi \neq \eta$.



Sources: Allen (1992), Crafts (1985), Wrigley and Schofield (1981)

Figure 1.5: Simulation and Data for England 1700-1850

(2001).[41] The larger aggregate capital share now implies a smaller ϵ.[42] Since capital is more important in aggregate production, it generates more externalities; ϵ thus has to be smaller to maintain the observed relationship between capital and TFP growth.[43] The simulation results shown in the third row of table 1.4 reveal that the larger manufacturing capital share leads to accelerated growth of output, population, and capital stock, compared to the baseline. Again, the difference is not very large. Because of its greater capital share, the manufacturing sector now profits more from aggregate capital accumulation, and TFP rises relative to agriculture. The relative price of p_M thus falls. Since $p_M = p_K$, the price of investment also falls. Consequently, a given investment ratio leads to more capital deepening. Faster capital accumulation, on the other hand, implies more rapid TFP growth, creating a virtuous circle.

Our final sensitivity check examines the elasticity of intertemporal substitution, $1/\psi$. The usual range for the CRRA parameter ψ is between 1 and 4. While we used $\psi = 1$ in the baseline, we now choose $\psi = 4$. The last row of table 1.4 shows that growth and structural change occur somewhat faster than

[41]To ensure comparability, we use her figure for the "capital-energy aggregate in the manufacturing sector".

[42]Since $\gamma_A > 0$, we use the same procedure as in the baseline calibration. We obtain $\epsilon = 0.88$ from $\widehat{\eta\epsilon} = 0.44$ and $\eta = 0.5$. Note that for $\overline{\gamma}_K = 1.17\%$ we now have $\gamma_{T,M} > \gamma_{T,A}$, which deviates somewhat from the historical record.

[43]We also need to re-calibrate initial TFP in agriculture and manufacturing as well as the initial capital stock. These are $A_{A,0} = 0.505$, $A_M = 0.325$, and $K_0 = 1.79$, respectively.

Table 1.4: Sensitivity Analysis

Changed Parameters	y_{1850}/y_{1700}	k_{1850}/k_{1700}	N_{1850}/N_{1700}	$N_{M,1850}/N_{1850}$
none [Baseline Model]	2.31	2.28	2.93	0.57
$\gamma_A = 0$, $\epsilon = 1.5$	1.99	1.98	2.40	0.50
$\phi = 0.2$, $\eta = 0.5$, $\epsilon = 0.88$	2.61	3.11	3.18	0.45
$\psi = 4$	2.40	2.33	3.17	0.57

in the baseline simulation. This might be considered counterintuitive. In one-sector growth models, the growth rate typically depends negatively on ψ. In a two-sector model, the relative price of manufacturing output can change. The baseline simulation has p_M rising initially, and then falling steadily.[44] In the baseline simulation with log-utility, the change in p_M has no impact since the income and substitution effect cancel each other. With $\psi > 1$, however, the income effect is relatively stronger. An (expected) increase in the price of consumption lowers today's expenditure and yields an increase in investment. Consequently, with $\psi = 4$ the investment rates at the beginning are larger than in the baseline simulation, which explains the faster growth.[45]

1.4.4 The Role of Chance

Adding shocks to our model produces a significant dispersion of industrialization outcomes. It also slows development on average. In the stochastic simulations, a negative shock lowers both total income and investment. Moreover, large negative shocks lead to starvation and a net decline of population and capital stock, reducing the scope for the capital externality to work its wonders. There is also a second, more subtle effect: In the stochastic simulation, a positive shock to agricultural productivity causes a surge in expenditure, and more demand for manufacturing goods. Investment increases. However, the positive shock to agricultural productivity also makes food much cheaper. This produces an increase in the relative price of capital so that a given quantity of savings translates into relatively less capital accumulation. By contrast, in the deterministic simulation, the relative price of capital (produced in manufacturing) does not change quickly, because agricultural and manufacturing TFP grow in tandem.

Figure 1.6 shows the results of 1,000 model runs, starting with the parameters for 1700, and simulating the model over 150 periods.[46] The share of the workforce in manufacturing is our indicator of industrialization. It varies substantially across simulations, and so does the growth rate.

The results lend support to Crafts's (1977) argument that historical accident may have contributed to England industrializing first – the range of outcomes is wide. Also, the actual historical performance of the English economy is in the better half of possible results. Most likely, England would have had markedly lower per capita income and experienced an even slower shift out of agriculture – many sim-

[44] The explanation is provided by equations (1.14) and (1.15). Initially, capital accumulation proceeds slowly, such that γ_K is small and $\gamma_{T,M} < \gamma_{T,A}$. Thus, the relative price of manufacturing increases. In later periods, when γ_K is larger, the opposite is true.

[45] Although investment rates are lower in later periods, the virtuous capital-externality-circle initiated in the early periods prevails.

[46] The model is solved numerically as described in Appendix A.6. For the stochastic simulations, we allow agricultural TFP A_A to follow the random process described in section 1.3.3.

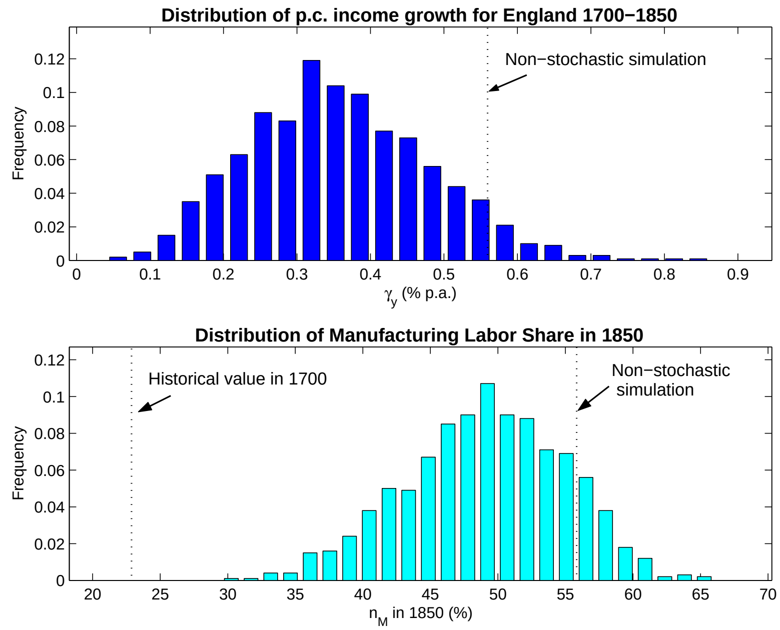


Figure 1.6: Stochastic Simulation for England 1700-1850

ulation values for 1850 are as much as one third lower. A run of good years in the 1740s aided the transformation, by producing higher incomes. This then led to higher demand for manufactured products [Gilboy (1932)]. To our knowledge, this is the first calibration exercise that demonstrates, based on a fully specified model, the extent to which Britain's industrial dominance in 1850 was the result of a lucky draw. Analogously, it could be argued that other, unmodelled factors – such as the Glorious Revolution's strengthening of property rights emphasized by North and Thomas (1973) – facilitated the acceleration of actual growth compared to the predicted rate. If chance could have played a role in the absolute rate of progress after 1700, it is natural to ask if it also played a role in determining which country got to be the First Industrial Nation. This is what we examine next.

1.4.5 Probabilities of Industrialization in other Countries

Why did England industrialize first? Could it have been France or China instead? In our model, industrialization occurs stochastically, but initial income, inequality, and the demographic regime are crucial determinants. Starting positions differed a good deal. We summarize some key variables in Table 1.5. England was both richer and more urbanized than France and China in 1700.

In order to compare England's chances of industrialization in 1700 with those of other countries, we need detailed, reliable data on per capita incomes, birth rates, and income support. There has recently been an upsurge of historical research on Chinese wages and the productivity of its agricultural sector. Revisionists' arguments along the lines of Pomeranz (2000) have proven to be overoptimistic about living standards in China. Land productivity was impressive, especially in the Yangtze delta. In some parts, where political pressures limited rent increases, peasants could live quite well, especially when measured in terms of the price of agricultural goods. Where silver wages are used – more relevant for our com-

Table 1.5: Income, Urbanization and Population Growth in other Countries

	p.c. income* (in 1990 Geary-Khamis dollars)		Population growth (% p.a.)		Urban Shares (%)		N_M/N** (%)
Year	1700	1820	1700-1820	1820-1850	1700	1800	1700
England	1250	1706	0.76***	0.83	13.3	20.3	23
France	910	1135	0.31	0.51	9.2	8.8	16
China	600	600	0.85	0.26	6.0	3.8	10

Sources: Maddison (2003) for p.c. income and population growth; Vries (1984) and Rozman (1973) for urban shares.
* Maddison figures are controversial. In the calibration we rely on urban share.
** Manufacturing Labor Share. For England: Calculated from Crafts (1985), leaving out services. For France and China: Author's calculation based on urban shares, assuming that the British ratio of the urban share to total employment in manufacturing is indicative of ratios elsewhere.
***1701-1751: 0.25%, from Wrigley and Schofield (1981)

parisons of the ability to purchase goods other than food – the gap between Europe and China is wide (Broadberry and Gupta 2005a). Without politically skewed incomes, however, conditions were much less favorable. Laborers were paid poorly. Everywhere after 1620, as a result of population pressure, "the downward trend toward immiseration is stark" (Allen 2006).[47] In the most comprehensive study of comparative living standards yet, Allen et al. (2005) found that Chinese families' incomes lagged behind Northern European ones by a large margin. Only in the impoverished South of Europe – like Milan – were living standard comparable. Allen et al. (2005) suggest that the ratio of English to Chinese wages may have been close to 2:1. Broadberry and Gupta (2005a) derive even more pessimistic figures for silver wages in the Yangtze.

In order to stack the odds in favor of China's prospects, we consistently make the most optimistic assumptions possible. We concentrate on the most prosperous area, the Yangtze. Instead of using the wage comparisons directly, we rely on the agricultural and non-agricultural labor shares in England, France and China, as based on urban shares. Table 1.6 gives our calibration figures. We calibrate TFP in agriculture and manufacturing such that we match both net population growth and manufacturing labor shares as close as possible.[48] Given the calibrated TFP, we can compute the implied consumption level relative to subsistence, $c/\underline{c}$. As Allen et al. (2005) argue, there is indeed a substantial part of the population in China that is not able to satisfy basic subsistence needs, as reflected by $c/\underline{c} < 1$. As given in equation (1.12), this corresponds to starvation of part of the population such that $\gamma_N < \gamma_b$. In France, on the other hand, $c > \underline{c}$ such that everyone survives and $\gamma_N = \gamma_b$. Note that since we avoided the pitfalls of size effects, TFP growth in our model is not dependent on size of the economy as such, but is driven by capital accumulation. More capital in agriculture yields the same benefits as in manufacturing. This biases our results towards industrialization in China, since there almost all capital is used in agriculture in 1700.

The birth rates for both France and China are constant. In the case of France, this is a simplification –

[47] Allen (2005a) does not provide figures for 1700, offering estimates for 1620 instead. His finding of a strong trend towards immiseration, and of broady comparable starting levels in incomes sustained by politically biased distribution of rents, is consistent with our argument here.

[48] Deviations from historical population growth and manufacturing labor shares as given in table 1.5 are small. As in the English case, manufacturing TFP is chosen such that $p_M = 2$.

Table 1.6: Calibration of Initial Conditions in Cross-Country Simulations

	$\frac{N_M}{N}$	γ_b	γ_N	$\frac{c}{\underline{c}}$
England	23%	0.29%	0.28%	1.101
France	16%	0.32%	0.32%	1.045
China*	11%	4.0%	0.73%	0.969

*Referring to the Yangtze Delta

population growth was low, and birth rates declined after 1800 in parallel with death rates. Wrigley and Schofield (1981) show that France had more of a "high pressure" demographic regime, with birth rates responding too little to avoid additional adjustment through the positive check. We deliberately simplify to highlight the importance of the demographic regime, and assume a constant birth rate to match observed population growth rates. For China, we also use historical data on population growth. We observe maximum fertility rates in the period immediately before the demographic transition in China [Chesnais (1992)], which implies birth rates of 4% in our setup (with infinitely-lived agents). Analogous to the British case, French shocks are derived from the movement of grain wages [Labrousse et al. (1970)].[49] For China, we used both the French and the British shock parameters in the stochastic simulation, but the results do not differ. According to our assumptions in section 1.3.9, there is also no redistribution to support lower incomes during times of crisis in France and China.[50]

For China, our simulations on average predict a decline in per capita income, combined with a very low labor share in manufacturing.[51] There are some cases of industrial development, but they are rare and stop far short of the extent of industrialization witnessed in England. The periods of benign development result from a sequence of positive shocks, which leads to capital accumulation outpacing demographic growth. As aggregate capital grows, the externality pushes up TFP. Eventually, the investment schedule crosses the line defined by $\delta + \gamma_N$ twice: from below for lower k and from above for higher k (due to decreasing returns to capital the investment schedule eventually becomes downward-sloping). The latter is a stable equilibrium with growing population and p.c. income .

France has markedly higher probabilities of industrializing than China. Its average share of the labor force in manufacturing in our simulations is 36.5 percent – much less than Britain, but a long way away from pre-industrial stagnation. Growth is markedly slower, at less than half the British rate. The two distributions overlap to some extent. As Crafts (1977) argued, much of the difference between the experience of France vs. England could be due to chance. Detailed examinations of "France's failure" may have suffered from hindsight bias, finding causes where there was simply bad luck. In comparison with China, on the other hand, chance plays almost no role – the British performance in all of our simulations is markedly better than the best possible one for China.

[49] We use figures for 1726-1792 and find $\theta = 0.595$ (t=5.71) and $\sigma_\varepsilon = 0.13$.
[50] As in the case of England, the effect of redistribution is negligible.
[51] This is in line with the tendency towards involution found by Allen (2005a).

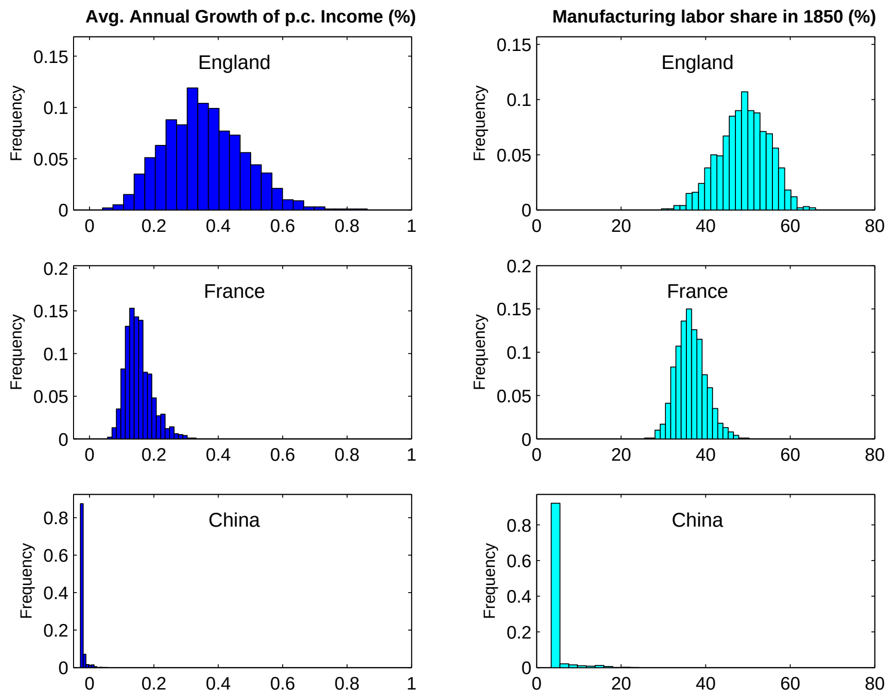


Figure 1.7: Stochastic Simulation for 1700-1850

1.4.6 Turning England into China

What was crucial about England's starting conditions – its demographic regime, its favorable income level, or the redistributive institutions that raised incomes for the bottom 40 percent of the population?

First, we simulate the development of the British economy using the same parameters as in the baseline calibration before, but changing the birth rate to a constant 4%. In the majority of simulations, after introducing the Chinese demographic regime in England, individual consumption declines to subsistence, so that all consumption expenditure goes to agriculture ($c_A/e = 1$). The economy typically starts near point C (figure 1.2). Population grows very quickly because higher initial incomes reduce death rates. Despite high savings rates, capital per head declines. This pushes the economy towards A. Instead of growing by 0.34% per year, per capita incomes now fall by 0.15% annually. As a result of high birth rates, N can grow quite quickly in a short period of time. Eventually, the economy reaches a stable equilibrium at point A, where during periods with average productivity, the only demand for manufactured goods comes from investment. Over the period as a whole, demographic growth will be slower than in the baseline case due to high mortality, driven by falling living standards. Our results therefore suggest that, instead of being able to industrialize, England would have seen a economic collapse with a high-pressure demographic regime. This underlines the crucial importance of fertility limitation as part of Europe's unique demographic regime. In passing, we might want to note that the parts of Europe where the European Marriage Pattern was weakest [Southern and Eastern Europe (Hajnal 1965)] also suffered from long delays before industrial development got under way.[52]

[52] Japan is also a case in point for our model – as noted by Mosk [1976], it had strikingly low fertility during the Tokugawa period.

What is the importance of starting conditions? Is a high starting point crucial for England's high chance to industrialize? We can repeat the simulations with Chinese starting conditions, but an English demographic regime. In figure 1.2 (left panel), the inv/k curve is shifted down and to the right. The economy at B will now grow more slowly as aggregate capital accumulation slows to a crawl. As table 1.7 shows, with Chinese starting levels and the English demographic regime, the country would have seen slow growth of per capita incomes. The share of the population in manufacturing rises gradually, eventually surpassing England's level in 1700. Population may stagnate or even fall because initially, many households are near the subsistence minimum.[53]

For our final counterfactual, we examine the effects of redistribution. In our model, the Poor Law is potentially important because it ensures that the malnourished can work even during years with poor harvests. We model this by assuming that in the absence of redistribution, during crisis periods ($c < \underline{c}$) the part of the population that will starve also does not work [reflecting the basic insight from Fogel (1994)]. Adding this effect to our simulations amplifies the impact of negative shocks in the short run. Over the long run, it hardly matters at all because higher land-labor ratios have a stabilizing influence. In China, the absence of redistribution makes catastrophic declines of population and output more likely. As noted by Lagerlöf (2006), these are a constant feature of the pre-industrial world. As it happens, British per capita incomes are too high even in 1700 for this mechanism to make much of a difference. Very good outcomes – showing growth above 0.6% p.a. – are more common in the simulations with redistribution, but the average is basically the same for stochastic simulations with and without the Poor Law.

Table 1.7: Counterfactual Simulations for Britain - Results for 1850

	p.c. Income Growth	Population Growth	Labor Share in M
Baseline Model	0.34%	0.57%	49.1%
Chinese Demography	-0.15%	0.12%	4.6%
Chinese Starting Levels	0.13%	-0.08%	28.2%
No Subsidies to the Poor	0.33%	0.56%	48.4%

Note: All results are the median of 1000 stochastic simulations, each over 150 periods.

1.5 Conclusions

This paper offers quantitative answers to our two initial questions: "why England?" and "why Europe?" Based on a calibrated two-sector growth model with an aggregate capital externality, we argue that Europe's unique demographic regime ensured starting positions that made industrial development much more likely. No lucky accident through a few good harvests, or as a result of natural resource endowments, could have similarly raised the chances to industrialize. Nor could redistribution, on its own, have had sufficiently benign effects.

We derive a model that focuses on the first transition in unified growth theory – from Malthusian stagnation to a post-Malthusian regime [Galor and Weil (2000), Galor (2005)]. The key driving variable is not the generation of ideas through a link with population size or an increase in the population's

Infanticide, not fertility limitation through changes in nuptiality, may have been decisive.

[53] Because of low fertility in the English demographic regime, recovery from negative shocks takes a long time.

quality. Factors highlighted by historians of technology play a crucial role – such as the importance of chance in new inventions, the role of tinkering, and the essentially non-economic motives for innovation [Mokyr (1990)]. All of this suggests that the biggest single determinant of technological progress was not the patent system, nor population size, but the fertility regime and the use of differentiated capital inputs. Interacting with the installed stock of machinery created the opportunities for "microinventions", in Mokyr's phrase.

England's chances to make many microinventions were good mainly because of high per capita incomes, resulting from fertility restriction. A more effective work force because of redistributive institutions raised output and increased industrialization probabilities, but this channel's role was small. These conclusions follow from our simulations of England's Industrial Revolution, which show a close fit between historical fact and model output. We also show how important it was that population growth accelerated in a context of high per capita incomes. Economic historians have long puzzled over the fact that the country with the biggest population increase between 1550 and 1800 also saw the biggest increase in per capita output [Wrigley (1988)]. In our model, this is no accident, but arises naturally from the interaction of starting conditions, the demographic regime, and the capital-use externality.

Based on the baseline simulation for England, we vary the parameter values to examine France's and China's chance to develop. The exercise suggests that France had reasonable prospects to develop, too. The absence of the Poor Law and a more high-pressure demographic regime reduced its chances, but not to such an extent that history could not have played itself out differently. The answer for China is fundamentally different. Because of the capital-diluting effects of rapid population growth, its chances of industrializing were very small. Only very unlikely sequences of good shocks could have given it a chance to develop. Starting conditions themselves were of secondary importance compared to the long-run influence of the demographic regime.

Our results also highlight one mechanism through which inequality in the early stages of development may be growth-reducing – if nutrient availability overall is low, redistribution from top to bottom may create opportunities for growth because it raises the workforce's effectiveness. This could qualify the conclusions by Galor and Moav (2004). They argued that greater inequality is beneficial when physical capital accumulation is key. The Galor and Moav effect may be conditional on overall nutrient supply being sufficiently generous to leave all groups of society in a position to perform hard labor. In our simulations, however, the consequences of workforce effectiveness matter, but are never large enough in the long run to dominate our results.

Economic historians have sometimes been sceptical that endogenous growth models can capture the complexity of the historical industrialization experience. Standard modelling approaches grappled with cross-sectional differences in timing and speed. Crafts (1995) concluded that the contrasting experiences of France and England did not seem to fit the mould of earlier models. Because of this, interpretations based on exogenous growth should be preferred. Our results demonstrate that more recent advances in unified growth theory can do much to resolve seeming contradictions between the historical record and growth models. In particular, the emphasis on capital accumulation and declining constraints on population growth during the first transition from stagnation to the post-Malthusian state prove useful. In this way, rigorous, quantitative examinations of the cross-sectional differences in the industrialization process can yield important conclusions about the nature of early development.

1.6 APPENDIX

1.6.1 Appendix A.1 – Optimization of Production

In this section of the appendix we derive the first order conditions (FOC) for profit-maximization of the production side of the model and calculate the demand function for capital varieties.

Final sector firms take input and output prices as given. A unit of capital variety j has value $p(j)$ and is borrowed at the gross interest rate R_K. Labor and land are paid wage w and land rental rate r_L, respectively. Final producers solve the following problems in agriculture and manufacturing production:

$$\max\{Y_A - \int_0^J R_K p(j)\nu_A(j)dj \; - wN_A - r_L L\} \tag{A.1}$$

$$\max\{p_M Y_M - \int_0^J R_K p(j)\nu_M(j)dj \; - wN_M\} \tag{A.2}$$

subject to the production functions (2.3) and (2.5). Capital producing firms take input prices as given but set the price of their own output in order to maximize profits. For given input prices, they solve the cost minimization problem

$$\min\{\int_0^J R_K p(j)\nu_{\tilde{j}}(j)dj \; + wN_{\tilde{j}} \; - \lambda_{\tilde{j}}\,[\nu(\tilde{j}) - \overline{\nu}(\tilde{j})]\} \tag{A.3}$$

subject to the production function (1.8), where $\overline{\nu}(\tilde{j})$ is the targeted production amount of variety $\tilde{j}$ and $\lambda_{\tilde{j}}$ is a Lagrange multiplier. In the following we derive the first order conditions for problems (A.1) - (A.3) and use them to obtain the demand function for capital varieties.

For agricultural output, equation (A.1) has the FOC

$$R_K p(j) \;=\; \phi\,\nu_A(j)^{-\frac{\epsilon}{1+\epsilon}} A_A \left[\int_0^J \nu_A(j)^{\frac{1}{1+\epsilon}} dj\right]^{\phi(1+\epsilon)-1} N_A^{\mu} L^{1-\phi-\mu},\ \forall j \tag{A.4}$$

$$w \;=\; \mu\, A_A \left[\int_0^J \nu_A(j)^{\frac{1}{1+\epsilon}} dj\right]^{\phi(1+\epsilon)} N_A^{\mu-1} L^{1-\phi-\mu} \tag{A.5}$$

$$r_L \;=\; (1-\phi-\mu)\, A_A \left[\int_0^J \nu_A(j)^{\frac{1}{1+\epsilon}} dj\right]^{\phi(1+\epsilon)} N_A^{\mu} L^{-\phi-\mu} \tag{A.6}$$

The corresponding FOC for manufacturing production follow from (A.2)

$$R_K p(j) \;=\; \eta\,\nu_M(j)^{-\frac{\epsilon}{1+\epsilon}} p_M A_M \left[\int_0^J \nu_M(j)^{\frac{1}{1+\epsilon}} dj\right]^{\eta(1+\epsilon)-1} N_M^{1-\eta},\ \forall j \tag{A.7}$$

$$w \;=\; (1-\eta)\, p_M A_M \left[\int_0^J \nu_M(j)^{\frac{1}{1+\epsilon}} dj\right]^{\eta(1+\epsilon)} N_M^{-\eta} \tag{A.8}$$

Finally, the cost-minimization problem (A.3) of a capital variety producer $\tilde{j}$ implies

$$R_K p(j) \;=\; \eta\,\nu_{\tilde{j}}(j)^{-\frac{\epsilon}{1+\epsilon}}\, \lambda_{\tilde{j}}\, A_{\tilde{j}} \left[\int_0^J \nu_{\tilde{j}}(j)^{\frac{1}{1+\epsilon}} dj\right]^{\eta(1+\epsilon)-1} N_{\tilde{j}}^{1-\eta},\ \forall j \tag{A.9}$$

$$w \;=\; (1-\eta)\, \lambda_{\tilde{j}}\, A_{\tilde{j}} \left[\int_0^J \nu_{\tilde{j}}(j)^{\frac{1}{1+\epsilon}} dj\right]^{\eta(1+\epsilon)} N_{\tilde{j}}^{-\eta} \tag{A.10}$$

Note that we have not imposed symmetry of capital variety prices in any of these derivatives. Rather, we will obtain symmetry in the following steps, which lead to the demand function for capital varieties. Equations (A.9) and (A.10) can be used to derive

$$\nu_{\widetilde{\jmath}}(j)^{\frac{1}{1+\epsilon}} = \left[\frac{w}{1-\eta} \frac{\eta}{R_K} \frac{1}{\int_0^J \nu_{\widetilde{\jmath}}(j)^{\frac{1}{1+\epsilon}} dj} N_{\widetilde{\jmath}} \frac{1}{p(j)} \right]^{\frac{1}{\epsilon}} \tag{A.11}$$

Integrating over all varieties $j \in [0, J]$ yields

$$\left[\int_0^J \nu_{\widetilde{\jmath}}(j)^{\frac{1}{1+\epsilon}} dj \right]^{1+\epsilon} = \frac{w}{1-\eta} \frac{\eta}{R_K P_J} N_{\widetilde{\jmath}} \tag{A.12}$$

where P_J is the price index of existing capital varieties $j \in [0, J]$, given by

$$P_J \equiv \left[\int_0^J p(j)^{-\frac{1}{\epsilon}} dj \right]^{-\epsilon} \tag{A.13}$$

We will need labor demand $N_{\widetilde{\jmath}}$ as a function of a given amount of output of variety $\widetilde{\jmath}$, $\overline{\nu}_{\widetilde{\jmath}}$, later on. To obtain this we plug (A.12) into the production function (1.8), which gives[54]

$$N_{\widetilde{\jmath}} = \frac{1-\eta}{w} \frac{1}{A_{\widetilde{\jmath}}} \left(\frac{R_K P_J}{\eta} \right)^{\eta} \left(\frac{w}{1-\eta} \right)^{1-\eta} (\overline{\nu}_{\widetilde{\jmath}} + F) \tag{A.14}$$

We then derive the demand for an existing variety j by a producer of a new variety $\widetilde{\jmath}$ by plugging (A.12) into (A.11) and substituting $N_{\widetilde{\jmath}}$ from (A.14)

$$\nu_{\widetilde{\jmath}}(j) = \frac{\eta}{R_K P_J} \left[\frac{P_J}{p(j)} \right]^{\frac{1+\epsilon}{\epsilon}} \frac{1}{A_{\widetilde{\jmath}}} \left(\frac{R_K P_J}{\eta} \right)^{\eta} \left(\frac{w}{1-\eta} \right)^{1-\eta} (\overline{\nu}_{\widetilde{\jmath}} + F) \tag{A.15}$$

Demand for variety j by a producer of a new variety depends on the price of j relative to the aggregate price index of capital varieties P_J. Note that $\overline{\nu}_{\widetilde{\jmath}}$ denotes the amount of the new variety $\widetilde{\jmath}$ that is actually produced, whereas $\nu_{\widetilde{\jmath}}(j)$ is the amount of an existing variety j used in the corresponding production process. We can now derive the total cost of producing $\overline{\nu}_{\widetilde{\jmath}}$ from (A.13) - (A.15):

$$C_{\widetilde{\jmath}} = \int_0^J R_K p(j) \nu_{\widetilde{\jmath}}(j) dj + w N_{\widetilde{\jmath}} = \frac{1}{A_{\widetilde{\jmath}}} \left(\frac{R_K P_J}{\eta} \right)^{\eta} \left(\frac{w}{1-\eta} \right)^{1-\eta} (\overline{\nu}_{\widetilde{\jmath}} + F) \tag{A.16}$$

Consequently, the marginal cost of variety $\widetilde{\jmath}$ production is given by

$$MC_{\widetilde{\jmath}} = \frac{1}{A_{\widetilde{\jmath}}} \left(\frac{R_K P_J}{\eta} \right)^{\eta} \left(\frac{w}{1-\eta} \right)^{1-\eta} \equiv MC_{\widetilde{J}}, \ \forall \widetilde{\jmath} \tag{A.17}$$

Marginal costs are the same for all capital variety producers $\widetilde{\jmath}$, which is one of the steps in our derivation of the symmetric equilibrium. We need two more ingredients to derive total demand for variety j, $\nu^d(j)$: the demand for j by agricultural and by manufacturing production. Using the FOC (A.4) - (A.8) and the production functions (2.3) and (2.5) we repeat the steps outlined in (A.11) - (A.15) and obtain

$$\nu_A(j) = \frac{\phi}{R_K P_J} \left[\frac{P_J}{p(j)} \right]^{\frac{1+\epsilon}{\epsilon}} \frac{1}{A_A} \left(\frac{R_K P_J}{\phi} \right)^{\phi} \left(\frac{w}{\mu} \right)^{\mu} \left(\frac{r_L}{1-\phi-\mu} \right)^{1-\phi-\mu} Y_A \tag{A.18}$$

$$\nu_M(j) = \frac{\eta}{R_K P_J} \left[\frac{P_J}{p(j)} \right]^{\frac{1+\epsilon}{\epsilon}} \frac{1}{A_M} \left(\frac{R_K P_J}{\eta} \right)^{\eta} \left(\frac{w}{1-\eta} \right)^{1-\eta} Y_M \tag{A.19}$$

[54]In this step we implicity impose that the constraint in (A.3) holds with equality, i.e., production is at its efficiency frontier.

Total demand for an existing intermediate variety j can be derived from (A.15), (A.18), and (A.19):

$$\nu^d(j) = \left[\frac{P_J}{p(j)}\right]^{\frac{1+\epsilon}{\epsilon}} \Phi \tag{A.20}$$

where

$$\begin{aligned}\Phi \quad \equiv \quad & \int_0^1 \frac{\phi}{R_K P_J}\frac{1}{A_A}\left(\frac{R_K P_J}{\phi}\right)^{\phi}\left(\frac{w}{\mu}\right)^{\mu}\left(\frac{r_L}{1-\phi-\mu}\right)^{1-\phi-\mu} Y_A(i)\,di\ + ... \\ & ... \int_0^1 \frac{\eta}{R_K P_J}\frac{1}{A_M}\left(\frac{R_K P_J}{\eta}\right)^{\eta}\left(\frac{w}{1-\eta}\right)^{1-\eta} Y_M(i)\,di\ + ... \\ & ... \int_0^{\widetilde{J}} \frac{\eta}{R_K P_J}\frac{1}{A_{\widetilde{\jmath}}}\left(\frac{R_K P_J}{\eta}\right)^{\eta}\left(\frac{w}{1-\eta}\right)^{1-\eta} (\overline{\nu}_{\widetilde{\jmath}} + F)\ d\widetilde{\jmath}\end{aligned} \tag{A.21}$$

The first two rows in (A.21) represent total demand for variety j from final producers $i \in [0,1]$ (i.e., from agriculture and manufacturing), and the last row is demand from currently active new variety producers $\widetilde{\jmath} \in [0, \widetilde{J}]$. Note that the price of variety j enters Φ only through the aggregate price index P_J, so that its effect on Φ is negligible. Consequently, Φ is treated as a constant in a capital variety producer's profit maximizing price decision:

$$\max_{p_{\widetilde{\jmath}}}\{p_{\widetilde{\jmath}}\nu^d(p_{\widetilde{\jmath}}) - C_{\widetilde{\jmath}}\left(\nu^d(p_{\widetilde{\jmath}})\right)\} \tag{A.22}$$

where $\nu^d(p_{\widetilde{\jmath}})$ is the total demand for the new capital variety $\widetilde{\jmath}$. Using (A.16), (A.17), and (A.20), we obtain the profit-maximizing price as a markup over marginal cost of production $MC_{\widetilde{\jmath}}$, which is the same for each capital variety producer, so that the price of all newly produced capital varieties in a given period is the same:

$$p_{\widetilde{\jmath}} = (1+\epsilon)MC_{\widetilde{J}} \equiv p_{\widetilde{J}},\ \forall\widetilde{\jmath} \tag{A.23}$$

Free entry into the capital producing sector implies that each firm $\widetilde{\jmath}$ makes zero profits, i.e., (A.22) is zero. This, together with the optimal price $p_{\widetilde{J}}$ from (A.23) implies

$$\nu(\widetilde{\jmath}) = \frac{F}{\epsilon},\ \ \forall\widetilde{\jmath} \tag{A.24}$$

That is, the amount of each newly produced capital variety, $\nu(\widetilde{\jmath})$, is the same in a given period, and moreover, is constant over time, even if factor prices and thus marginal costs change.

1.6.2 Appendix A.2 – Capital Varieties and Aggregate Capital

In the following, we refer to aggregate capital as the collection of all machines available for production in a given period:

$$K = \int_0^J \nu(j)dj \tag{A.25}$$

where $\nu(j)$ is the amount of capital variety j when it was produced ($\nu(j)$ does not change until j depreciates – it then becomes zero). We choose the fixed cost F such that $F = \epsilon$. Equation (A.24), and the fact that $\nu(j)$ is constant, imply:

$$\nu(\widetilde{\jmath}) = \nu(j) = 1,\ \ \forall\widetilde{\jmath}, j \tag{A.26}$$

Therefore, the amount of each capital variety circulating in the economy (new and existing ones) is the same. Our choice of F serves to simplify the following analysis since it implies, together with (A.25), that

$$J = K \tag{A.27}$$

that is, the total amount of capital in the economy is equal to the amount of capital varieties. Moreover, newly produced capital is given by $\int_0^{\widetilde{J}} \nu(\widetilde{\jmath})d\widetilde{\jmath} = \widetilde{J}$. Consequently, $\widetilde{J}_t$ denotes the mass of capital variety producers as well as the number of varieties that are produced in period t (but are used for production only from the next period on). The law of motion for the aggregate capital stock is thus equivalent to the one for varieties, $J_{t+1} = (1-\delta)J_t + \widetilde{J}_t$. The mass of currently active capital variety producers can be derived from total investment, $I = Y - eN$:

$$\widetilde{J}_t = \frac{I_t}{p_{\widetilde{J},t}} = J_{t+1} - (1-\delta)J_t \tag{A.28}$$

According to equation (A.23), the price of all newly produced capital varieties could differ from old varieties if marginal costs vary. We therefore add the assumption that owners of existing capital varieties exert the same market power as producers of new ones. All capital of variety j is owned by one individual or entity (although, of course, different entities can own different varieties).[55] The owner of an existing variety j chooses p_j to maximize $p_j\nu^d(p_j)$ subject to $\nu^d \leq 1$, since the amount owned of each j is one. Equation (A.20) with $\epsilon > 0$ implies that revenue $p_j\nu^d(p_j)$ is decreasing in the price of j. Therefore, owners of existing varieties want to charge the smallest possible price at which the constraint $\nu^d \leq 1$ holds. The constraint holds with equality if $p_j = p_{\widetilde{J}}$. Intuitively, if the owner of an existing capital variety chooses a price above $p_{\widetilde{J}}$, demand is lower than unity and part of the variety is wasted. This is not optimal because a marginal price decrease would raise the revenue and thus profits. On the other hand, if $p_j < p_{\widetilde{J}}$, demand is larger than unity and the fixed supply of one unit is not sufficient to satisfy demand. Thus, the price of existing and new capital varieties is the same within each period. We can now define the price of capital p_K:

$$p(\widetilde{\jmath}) = p(j) \equiv p_K, \ \ \forall \widetilde{\jmath}, j \tag{A.29}$$

Equation (A.26) establishes symmetry in capital producing sectors. In the following we slightly abuse notation and use $\widetilde{J}$ as the subscript for a representative new capital producer as well as for the mass of all producers of new capital varieties. Because the mass of final sector firms is one, output (Y_A, Y_M) and factor inputs (N_i, L, and $\nu_i(j)$ for $i = A, M$) of a representative final producer are equal to aggregate final output and inputs. The price equality of capital varieties j given in equation (A.29), used in (A.15), (A.18), and (A.19), implies that firms use the same amount of each variety, i.e., $\nu_i(j) = \nu_i$, $\forall j$ and $i = A, M, \widetilde{J}$. Clearly, the total amount demanded of each variety (i.e., the integral of $\nu_i(j)$ over all producers i) is also equal for all j: $\nu(j) = \nu$. Market clearing of each existing variety j then requires[56]

$$\nu_A + \nu_M + \widetilde{J}\nu_{\widetilde{J}} = \nu = 1 \tag{A.30}$$

[55] We noted before that existing capital varieties $j \in [0, J]$ are owned by consumers. Our assumption thus requires that population N be a multiple of the measure of capital varieties J. To circumvent this problem we can assume that single consumers bring their money to banks and that these act as profit-maximizing owners of each capital variety.

[56] Recall that newly produced varieties are only used from the next period on, but existing varieties are used by the mass $\widetilde{J}$ of new varieties producers.

where the last equality follows from (A.26). At the aggregate level all capital, labor, and land are used in each period. Integrating (A.30) over all existing varieties $j \in [0, J]$ yields

$$J\nu_A + J\nu_M + J\,\widetilde{J}\nu_{\widetilde{J}} = J\nu = J = K \tag{A.31}$$

Recalling that $J = K$, we can interpret ν_A, ν_M, and $\widetilde{J}\nu_{\widetilde{J}}$ as the aggregate capital shares in agriculture, manufacturing and variety production, respectively.

1.6.3 Appendix A.3 – Market Clearing and Equilibrium

The market clearing conditions for single capital varieties and aggregate capital are given by (A.30) and (A.31), respectively. The corresponding conditions for labor and land are:

$$N_A + N_M + \widetilde{J}N_{\widetilde{J}} = N \tag{A.32}$$

$$\int_0^1 L di = L \tag{A.33}$$

where the latter condition is trivial because land is only used for agriculture by the $[0, 1]$ final sector firms. Market clearing in final product markets requires:

$$Nc_A = Y_A \tag{A.34}$$

$$Nc_M = Y_M \tag{A.35}$$

Before defining the equilibrium, we need to introduce total nominal output Y, since this is the basis for individual income $y = Y/N$ that enters in consumers' intertemporal optimization decision. Let

$$Y = Y_A + p_M Y_M + p_K \widetilde{J} \tag{A.36}$$

where the last term represents the total value of newly produced capital varieties ($\widetilde{J}\nu$, with $\nu = 1$). This equation, together with (A.28), $J = K$, and the condition that consumers' budget constraints hold with equality ($Y_A + p_M Y_M = eN$), implies the law of motion for capital

$$K_{t+1} = (1 - \delta)K_t + (1/p_{K,t})(Y_t - e_t N_t) \tag{A.37}$$

that is taken into account in the intertemporal optimization (1.4) by households.[57]

Definition 1 *Given initial values $A_{A,0}$, N_0, $K_0 = \int_0^{J_0} \nu_0(j)dj = J_0$ (since $\nu_0(j) = 1$), and L, a competitive equilibrium consists of sequences for $t \geq 0$ of agricultural TFP, $\{A_{A,t}\}$; prices, $\{p_{M,t}, p_{K,t}, R_{K,t}, r_{L,t}, w_t\}$; final sector firm allocations $\{Y_{A,t}, Y_{M,t}, \nu_{A,t}, \nu_{M,t}, N_{A,t}, N_{M,t}, L_t\}$, capital sector firm allocations $\{\nu_t, N_{\widetilde{j},t}, \nu_{\widetilde{j},t}\}$ for all $\widetilde{j} \in [0, \widetilde{J}_t]$ producing at t; and household allocations $\{c_{A,t}, c_{M,t}\}$ such that (i) Given the sequence of prices, final sector firm allocations solve the problems specified in (A.1) and (A.2), and capital sector firm allocations solve (A.3); (ii) Producers of new capital varieties charge the profit-maximizing price given by (A.23), and, due to free entry, sell the amount given in (A.24) of each variety; (iii) Owners of existing capital varieties charge the price given by (A.29); (iv) Given the sequence of prices, consumer allocations maximize (2.1) subject to $c_{A,t} + p_{M,t}c_{M,t} \leq e_t$, and consumer consumption expenditures e_t satisfy the Euler equation (1.5); (v) The market clearing conditions (A.30)-(A.35) hold; (vi) The law of motion of capital is given by (A.37); and (vii) Population growth follows (1.11)-(1.13).*

[57] The term $(1 + \gamma_{N,t})k_{t+1}$ in (1.4) results from the fact that capital will be divided among $(1 + \gamma_{N,t})N_t$ household members in the next period.

1.6.4 Appendix A.4 – Aggregate Externality Representation

In this section we utilize the symmetry of capital variety use in production to derive a simplified representation of the model. It offers two advantages: First, the influence of aggregate externalities on productivity can easily be represented in the production functions. Second, with a single assumption about TFP in capital variety production, we simplify the model such that variety production can be included in the manufacturing sector. This reduces the number of equations that must be simulated to solve the model numerically.

Using $\nu_i(j) = \nu_i, \forall j$ and $i = A, M, \widetilde{J}$ in the production functions (2.3), (2.5), and (1.8), the integral over all capital varieties $j \in [0, J]$ simplifies to $J^{\phi\epsilon}(J\nu_A)^\phi$ in agriculture, $J^{\eta\epsilon}(J\nu_M)^\eta$ in manufacturing, and $J^{\eta\epsilon}(J\nu_{\widetilde{J}})^\eta$ in capital variety production. The terms in parentheses $(J\nu_i)$ represent the total capital used in firm i (i.e., the number of capital varieties multiplied by the amount utilized of each variety). It is convenient to simplify notation and label these terms $K_A \equiv J\nu_A$, $K_M \equiv J\nu_M$, and $K_{\widetilde{J}} \equiv J\nu_{\widetilde{J}}$. We also use (A.27) and set $J = K$. This implies the following simplified production functions:

$$Y_A = \quad A_A K^{\phi\epsilon} K_A^\phi N_A^\mu L^{1-\phi-\mu} \tag{A.38}$$

$$Y_M = \quad A_M K^{\eta\epsilon} K_M^\eta N_M^{1-\eta} \tag{A.39}$$

$$\nu(\widetilde{\jmath}\,) = \quad A_{\widetilde{J}} K^{\eta\epsilon} K_{\widetilde{J}}^\eta N_{\widetilde{J}}^{1-\eta} - F \quad = \nu, \ \ \forall\widetilde{\jmath} \tag{A.40}$$

where $N_{\widetilde{\jmath}} = N_{\widetilde{J}}, \ \forall\widetilde{\jmath}$ follows from symmetry in variety production and (A.14).

In the following steps we will derive the TFP paramater $A_{\widetilde{J}}$ as a function of A_M such that, despite the fixed cost in capital variety production (A.40), this sector's output can be described by the manufacturing production function (A.39). First, recall from (A.24) that each capital variety producer's output is F/ϵ. Second, we derive the labor and capital variety input needed to produce F/ϵ units of manufacturing output. Repeating steps (A.11) - (A.15) for the manufacturing sector and using symmetry of variety input prices yields

$$N_M \quad = \quad \frac{1-\eta}{w} \frac{1}{A_M J^{\eta\epsilon}} \left(\frac{R_K p_K}{\eta}\right)^\eta \left(\frac{w}{1-\eta}\right)^{1-\eta} Y_M \tag{A.41}$$

$$K_M \quad = \quad J\nu_M = \frac{\eta}{R_K p_K} \frac{1}{A_M J^{\eta\epsilon}} \left(\frac{R_K p_K}{\eta}\right)^\eta \left(\frac{w}{1-\eta}\right)^{1-\eta} Y_M \tag{A.42}$$

Third, we suppose that we want to produce F/ϵ units of a new capital variety $\widetilde{\jmath}$ – that is, using the capital variety technology (A.40) – with the labor and capital input given in (A.41) and (A.42) – i.e., the inputs needed when applying the manufacturing technology.

$$F/\epsilon \overset{!}{=} \nu(\widetilde{\jmath}\,) = A_{\widetilde{J}} K^{\eta\epsilon} \left(K_M|_{Y_M=F/\epsilon}\right)^\eta \left(N_M|_{Y_M=F/\epsilon}\right)^{1-\eta} - F \tag{A.43}$$

We now use the corresponding inputs, i.e., (A.41) and (A.42) evaluated at $Y_M = F/\epsilon$. This yields a constraint on the ratio of $A_{\widetilde{J}}$ and A_M:

$$A_{\widetilde{J}} = (1+\epsilon) A_M \tag{A.44}$$

The price of capital varieties, as implied by (A.23), (A.17), and price symmetry, is then

$$p_{\widetilde{J}} = (1+\epsilon) \frac{1}{A_{\widetilde{J}} K^{\eta\epsilon}} \left(\frac{R_K p_K}{\eta}\right)^\eta \left(\frac{w}{1-\eta}\right)^{1-\eta} = \frac{1}{A_M K^{\eta\epsilon}} \left(\frac{R_K p_K}{\eta}\right)^\eta \left(\frac{w}{1-\eta}\right)^{1-\eta} \tag{A.45}$$

which is equal to marginal cost of manufacturing production, as can be verified by calculating total cost $C_M = wN_M + R_K K_M$ from (A.41) and (A.42) and deriving it with respect to Y_M. Due to perfect competition in final production, the price of output equals marginal cost, i.e., $p_M = MC_M$. Consequently, (A.45) implies $p_{\tilde{\jmath}} = p_M$, and, using (A.29) we obtain

$$p_K = p_M \tag{A.46}$$

By choosing $A_{\tilde{\jmath}}$ according to (A.44), each capital variety producer uses exactly the amounts of labor and capital inputs that a manufacturing firm would need in order to produce the same (fixed) output F/ϵ and charges the same price that a manufacturing producer would request. Intuitively, this result follows because the higher TFP in variety production exactly offsets the fixed cost F. We can thus incorporate the capital variety producing sector in the manufacturing sector. The simplification follows independent of our assumption $F = \epsilon$ that leads to (A.26). Also, increasing returns in variety production imply that the TFP advantage $A_{\tilde{\jmath}} > A_M$, necessary to compensate for F, decreases with output $\nu(\tilde{\jmath})$. It is therefore crucial that each capital variety firm produces a constant amount of output, as follows from (A.24), such that the necessary TFP difference is the same for all variety producers and constant over time.

In addition to consumers' demand, Y_M^d, manufacturing must also satisfy the demand for capital investment, as given by (A.28).[58] Imposing market clearing, the total amount of manufacturing supply, Y_M, must thus equal demand from households and capital investment (I/p_K), where we can use $p_K = p_M$:

$$Y_M = Y_M^d + \frac{I}{p_M} \tag{A.47}$$

The simplified model is thus equivalent to a two-sector model where capital investment goods are produced in the manufacturing sector.

1.6.5 Appendix A.5 – Equilibrium Conditions of the 2-Sector Model

Having derived the two-sector version of the model in the previous section, we now present the corresponding equilibrium conditions. The FOC for agriculture and manufacturing profit maximization (A.4) - (A.8) can be easily simplified to their symmetric version by using $\nu_i(j) = \nu_i, \forall j$ and $i = A, M, \widetilde{J}$. The FOC's are the standard ones corresponding to profit maximization of (A.38) and (A.39). Factor payments to capital and labor are equal in both sectors, while land rents are determined in agriculture:

$$r_L = (1-\phi-\mu)A_A K^{\phi\epsilon} K_A^{\phi} N_A^{\mu} L^{-\phi-\mu} \tag{A.48}$$

$$p_K R_K = \phi A_A K^{\phi\epsilon} K_A^{\phi-1} N_A^{\mu} L^{1-\phi-\mu} = \eta p_M A_M K^{\eta\epsilon} K_M^{\eta-1} N_M^{1-\eta} \tag{A.49}$$

$$w = \mu A_A K^{\phi\epsilon} K_A^{\phi} N_A^{\mu-1} L^{1-\phi-\mu} = (1-\eta) p_M A_M K^{\eta\epsilon} K_M^{\eta} N_M^{-\eta} \tag{A.50}$$

Aggregate capital K and population N are given at the beginning of a period. In the following, we take per-capita expenditure e as given and solve for the intratemporal equilibrium. From this solution we obtain p_M and R_K, which we then use to solve the Euler equation (1.5). Total demand for agriculture products can be derived from (1.2): $Y_A^d = Nc_A = N\left[\alpha e + (1-\alpha)\underline{c}\right]$. The remaining expenditure goes to manufacturing, which implies $Y_M^d = Nc_M = N\left[(1-\alpha)(e-\underline{c})/p_M\right]$. Total demand for manufacturing

[58] Recall that the new capital produced in a given period is equal to $\int_0^{\widetilde{J}} \nu d\tilde{\jmath} = \widetilde{J}$ since $\nu = 1, \forall \tilde{\jmath}$.

is given by (A.47). Markets clear for agriculture and manufacturing output:

$$N\left[\alpha e+(1-\alpha)\underline{c}\right] = A_A K^{\phi\epsilon} K_A^{\phi} N_A^{\mu} L^{1-\phi-\mu} = Y_A \tag{A.51}$$

$$N\,\frac{(1-\alpha)}{p_M}\,(e-\underline{c})+\frac{I}{p_M} = A_M K^{\eta\epsilon} K_M^{\eta} N_M^{1-\eta} = Y_M \tag{A.52}$$

All land L is used and factor markets clear:

$$N_A+N_M = N \tag{A.53}$$

$$K_A+K_M = K \tag{A.54}$$

Finally, total nominal output, as given in equation (A.36), now simplifies to

$$Y=Y_A+p_M Y_M \tag{A.55}$$

This gives us a system with 13 unknowns: Y, Y_A, Y_M, I, N_A, N_M, ,K_A, K_M, R_K, r_L, w, p_M, p_K; and 13 equations: (A.46), (A.48), (A.53) - (A.55), and – each of the following counting twice – (A.49) - (A.52). Population growth is then derived from (1.12) and (1.13), where $c=c_A+c_M$ with $c_A=Y_A/N$ and $c_M=(p_M Y_M-I)/(p_M N)$. This system of equations characterizes the intratemporal equilibrium. We use the corresponding solution to derive per-capita expenditures e from the Euler equation in the iterative process described in the next section.

1.6.6 Appendix A.6 – Numerical Simulations

In this section we outline the simulation of the equilibrium given in the previous section. Dividing (A.49) by (A.50) and using (A.53) and (A.54) to substitute N_M and K_M implies

$$\frac{\phi}{\mu}\frac{N_A}{K_A}=\frac{\eta}{1-\eta}\frac{N-N_A}{K-K_A} \tag{A.56}$$

This condition, together with (A.51), gives a system of two equations with two unknowns, N_A and K_A, that we solve numerically for given e. Given N_A and K_A, the remaining variables can be derived analytically.

To solve the Euler equation (1.5) we use policy function iteration where expenditure is a linear function of the (given) per-capita capital at the beginning of a period:

$$e=\varphi k \tag{A.57}$$

We discretize the shocks ε to agricultural productivity using Gaussian quadrature with Q nodes and corresponding weights ω_q, defined by $\{\varepsilon_q,\omega_q\}_{q=1}^{Q}$. We use a projection method to solve for the coefficient φ, as described in the following steps:

1. Initialize by a guess φ_0 (a small positive number)

2. In iteration l, for φ_l, calculate e according to (A.57)

3. For the given e, obtain $y=Y/N$, γ_N, and $p_M=p_K$ from the the intratemporal equilibrium and calculate next period's population $N'=gN$

4. Evaluate the implied next period's p.c. capital, k', by

$$k' = \frac{1}{1+\gamma_N}\left[(1-\delta)k + \frac{y-e}{p_K}\right]$$

and calculate $K' = k'N'$.

5. Evaluate next period's consumption expenditure $e' = \varphi k'$. At all Gaussian quadrature nodes q, calculate $A'_{A,q} = z'_q \underline{A}'_A$, where $\underline{A}'_A = (1+\gamma_A)\,\underline{A}_A$ and $\ln z'_q = \theta \ln z + \varepsilon_q$

6. For the given N', K', e', and $A'_{A,q}$ obtain $y'_q = Y'_q/N'$, g'_q, and $p'_{M,q} = p'_{K,q}$ from the intratemporal equilibrium for $q = 1, ..., Q$

7. Evaluate expenditure implied by the Euler equation as

$$\widetilde{e} = \left[\beta \sum_{q=1}^{Q} \omega_q \left(\frac{p'_{K,q}}{p_K}\right)\left(\frac{p'_{M,q}}{p_M}\right)^{(1-\alpha)(\psi-1)} \left(\frac{1}{e'_q - \underline{c}}\right)^{\psi} \left(R'_{K,q} + 1 - \delta\right)\right]^{-\frac{1}{\psi}} + \underline{c}$$

8. Calculate $\widetilde{\varphi} = \widetilde{e}/k$

9. If $\|\varphi - \widetilde{\varphi}\| < 10^{-9}$, stop the iteration and accept φ as a solution. Otherwise use a Broyden solver to update φ_{l+1} and go to step 2. Repeat until convergence.

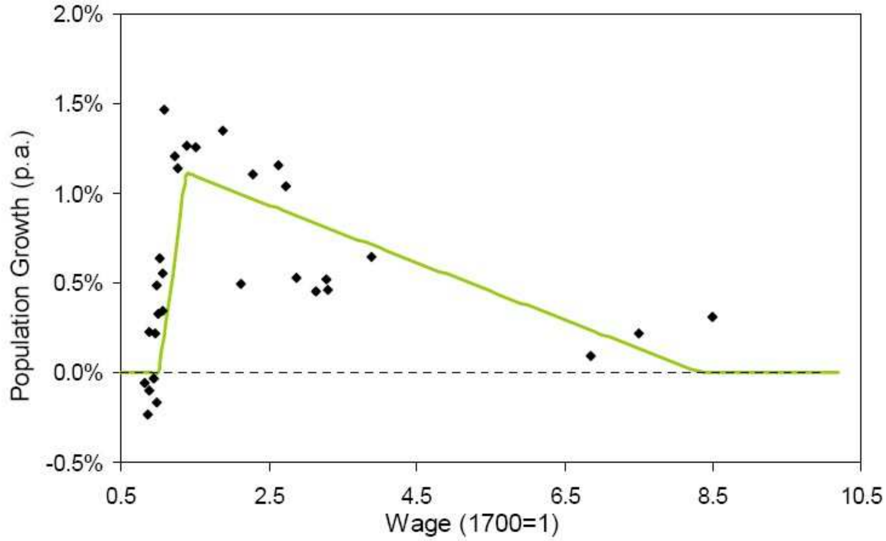
1.6.7 Appendix A.7 – Calibration of the British Birth Function

In this section we describe the calibration of the net birth rate function $g_b(c)$ based on British historical data, as shown in figure 1.8. Crucially for our purposes, English birth rates responded positively to higher wages, as demonstrated by Wrigley and Schofield (1981), and Wrigley et al (1997). To derive the function's exact shape, we use an exercise similar in spirit to Hansen and Prescott (2002). We employ a spline regression, defining $x = w/w_0$, where w_0 represents the wage in 1700. Population growth, is $y_p(x_t) = N_{t+1}/N_t - 1$. Let x_{peak} denote the cutoff-point at which the slope changes its sign from positive to negative. We then define a dummy $d = 1$, whenever $x \le x_{peak}$ and zero, else. Population growth in 1700 was close to zero; we thus impose $y_p(w_0/w_0) = y_p(1) = 0$. The spline regression is

$$y_p(x) = \beta_1 \underbrace{[(x-1)d + (x_{peak}-1)(1-d)]}_{x_1} + \beta_2 \underbrace{(x - x_{peak})(1-d)}_{x_2} + u \qquad \text{(A.58)}$$

where u is an error term. When running this regression (without constant) we choose the cutoff-point x_{peak} to maximize R^2, and obtain $\widehat{\beta}_1 = 0.0277$ (t=8.03) and $\widehat{\beta}_2 = -0.0016$ (t=-3.39); $x_{peak} = 1.4$; the adjusted R^2 is 0.70. As in Hansen and Prescott (2002), we impose that demographic growth rates cannot be negative because incomes are too high. For very low income, the net birth rate is zero and population diminishes due to starvation as described in section 1.3.10. The $g_b(\cdot)$ function for England is then defined as

$$g_b(x) = \begin{cases} \max\{\widehat{\beta}_1(x-1),\ 0\} + 1, & if\ \ x \le x_{peak} \\ \max\{\widehat{\beta}_1(x_{peak}-1) + \widehat{\beta}_2(x - x_{peak}),\ 0\} + 1, & else \end{cases} \qquad \text{(A.59)}$$



Sources: Population: Wrigley and Schofield (1981) for 1541-1871; Mitchell and Jones (1971) for 1881-1961. Real wages: Clark (2005)

Figure 1.8: Population Growth and Wage in England

References

[1] Allen, Robert C. (1992), "Enclosure and the Yeoman/the Agricultural Development of the South Midlands, 1450-1850." Oxford: OUP.

[2] Allen, Robert C. (2005a). "Real Wages in Europe and Asia: A First Look at the Long-term Patterns." In: Robert C. Allen, Tommy Bengtsson, and Martin Dribe, eds., Living Standards in the Past, Oxford: OUP.

[3] Allen, Robert C. (2005b). "Capital Accumulation, Technological Change, and the Distribution of Income during the British Industrial Revolution", Nuffield College working paper.

[4] Allen, Robert C. (2006). "Agricultural Productivity and Rural Incomes in England and the Yangtze Delta, c. 1620- c. 1820." Nuffield College working paper.

[5] Allen, Robert C., Jean-Pascal Bassino, Debin Ma, Christine Moll-Murata and Jan Luiten van Zanden (2005). "Wages, Prices, and Living Standards in China, Japan, and Europe, 1738-1925." Working paper.

[6] Antras, Pol and Hans-Joachim Voth (2003). "Factor Prices and Productivity Growth during the British Industrial Revolution." Explorations in Economic History 38: 52-77.

[7] Boyer, George (1990). "An Economic History of the English Poor Law, 1750-1850." Cambridge: CUP.

[8] Broadberry, Stephen and Bishnupriya Gupta (2005a). "The Early Modern Great Divergence: Wages, Prices and Economic Development in Europe and Asia, 1500-1800." Warwick working paper.

[9] Broadberry, Stephen and Bishnupriya Gupta (2005b). "Cotton Textiles and the Great Divergence: Lancashire, India and Shifting Competitive Advantage, 1600-1850." Warwick working paper.

[10] Braudel, Fernand (1973). "Capitalism and Material Life, 1400-1800." New York, Harper and Row.

[11] Cervellati, Matteo and Uwe Sunde (2005). "Human Capital Formation, Life Expectancy, and the Process of Development." American Economic Review 95(5): 1653-1672.

[12] Chesnais, Francois (1992). "The Demographic Transition." Oxford: OUP.

[13] CIA (2005). CIA fact book 2005: http://www.cia.gov/cia/publications/factbook/

[14] Clark, Gregory (2003a). "The Great Escape: The Industrial Revolution in Theory and in History." Unpublished manuscript, UC Davis.

[15] Clark, Gregory (2003b). "Agricultural Productivity, Prices and Wages." In Joel Mokyr (ed.) Encyclopedia of Economic History (Oxford: Oxford University Press, Vol 1: 92-96.

[16] Clark, Gregory (2005). "The Condition of the Working-Class in England, 1209-2004." Journal of Political Economy 113(6): 1307-1340.

[17] Cole, W.A. (1973). "Eighteenth-century Economic Growth Revisited." Explorations in Economic History 10: 327-48.

[18] Crafts, N.F.R. (1977). "Industrial Revolution in England and France: Some Thoughts on the Question, 'Why Was England First?'." Economic History Review, Second Series 30(3): 429-41.

[19] Crafts, N.F.R. (1985). "British Economic Growth During the Industrial Revolution." Oxford, OUP.

[20] Crafts, N.F.R. (1995). "Exogenous or Endogenous Growth? The Industrial Revolution Reconsidered." Journal of Economic History 55(4): 745-72.

[21] Crafts, N.F.R. and Knick Harley (1992). "Output Growth and the British Industrial Revolution: A Restatement of the Crafts-Harley View." Economic History Review 45: 703-730.

[22] Crafts, N.F.R. and Knick Harley (2000). "Simulating the Two Views of the Industrial Revolution." Journal of Economic History 60: 819-841.

[23] Diamond, Jared (2004). "Collapse: How Societies Choose to Fail or Succeed." New York: Viking.

[24] Feinstein, Charles (1981). "Capital Accumulation and the Industrial Revolution." In: R. Floud and D. McCloskey, eds., The Economic History of Britain since 1700, vol. I. Cambridge: CUP.

[25] Feinstein, Charles H. (1988). “The Rise and Fall of the Williamson Curve.” Journal of Economic History 48: 699-729.

[26] Fogel, Robert (1994). "Economic Growth, Population Theory, and Physiology: The Bearing of Long-Term Processes on the Making of Economic Policy." American Economic Review 84(3): 369-95

[27] Galor, Oded (2005). "From Stagnation to Growth: Unified Growth Theory." In: P. Aghion, S. Durlauf, eds., Handbook of Economic Growth, Vol. 1A. Amsterdam: North-Holland.

[28] Galor, Oded and Omer Moav (2002). "Natural Selection and the Origin of Economic Growth." Quarterly Journal of Economics 117(4): 1133-91.

[29] Galor, Oded and Omer Moav (2004). "From Physical to Human Capital Accumulation: Inequality and the Process of Development." Review of Economic Studies 71: 1001-1026.

[30] Galor, Oded and Andrew Mountford (2003). "Trade, Demographic Transition, and the Great Divergence: Why are a Third of People Indian or Chinese?" Hebrew University Working Paper.

[31] Galor, Oded and D. Weil (2000). "Population, Technology and Growth: From the Malthusian Regime to the Demographic Transition and Beyond." American Economic Review 89: 806-828.

[32] Gilboy, Elisabeth (1932). "Demand as a Factor in the Industrial Revolution." reprinted in: A.H. Cole, ed., Facts and Factors in Economic History. Cambridge, MA, Harvard University Press.

[33] Goldstone, Jack (2002). "Efflorescences and Economic Growth in World History:

[34] Rethinking the "Rise of the West" and the Industrial Revolution", Journal of World History 13(2): 323-389.

[35] Hajnal, John (1965). "European marriage in perspective". In: Glass, D. V. and Eversley, D. E. C. Population in History. London: Arnold.

[36] Hansen, Gary D. and Edward Prescott (2002). "Malthus to Solow." American Economic Review 92(4): 1205-1217.

[37] Horrell, Sara (1996). "Home Demand and British Industrialization." Journal of Economic History 56(3): 561-604.

[38] Jones, Charles (2001). "Was an Industrial Revolution Inevitable? Economic Growth over the Very Long Run." Advances in Macroeconomics 1(2).

[39] Kelly, Morgan (2005). "Living Standards and Population Growth: Malthus was Right". UC Dublin manuscript.

[40] King, Peter (1997). "Pauper Inventories and the Material Lives of the Poor in the Eighteenth and Early Nineteenth Centuries", in T. Hitchcock, P. King, and P. Sharpe, Chronicling Poverty: The Voices and Strategies of the English Poor, Basingstoke: Houndsmills.

[41] Kögel, Tomas, and Alexia Prskawetz (2001), "Agricultural Procutivity Growth and Escape from the Malthusian Trap", Journal of Economic Growth, 6: 337-357.

[42] Labrousse, Ernest, Ruggiero Romano and F.-G. Dreyfus (1970). "Le Prix Du Froment en France 1726-1913." Paris: SEVPEN.

[43] Lagerlöf, Nils-Petter (2003). "From Malthus to Modern Growth: Can Epedemics explain the three Regimes." International Economic Review 44(2): 755-777.

[44] Lagerlöf, Nils-Petter (2006). "The Galor-Weil Model Revisited: A Quantitative Exploration." Journal of Economic Dynamics 9(1): 116-142.

[45] Landes, David (1999). "The Wealth and Poverty of Nations: Why Some Are So Rich and Some So Poor." Harvard: Harvard University Press.

[46] Lindert, Peter H. (2000). "When Did Inequality Rise in Britain and America?" Journal of Income Distribution 9(1): 11-25.

[47] Lindert, Peter H. and Jeffrey G. Williamson (1982). "Revising England's Social Tables, 1688-1812." Explorations in Economic History 19(4): 385-408.

[48] Maddison, Angus (2003). "The World Economy: Historical Statistics." Paris, France, Development Centre of the OECD.

[49] Matsuyama, Kiminori (1991). "Increasing Returns, Industrialization, and Indeterminacy of Equilibrium." Quarterly Journal of Economics 106(2): 617-650.

[50] Matsuyama, Kiminori (2002). "The rise of mass consumption societies." Journal of Political Economy 110(5): 1035-1070.

[51] McCloskey, Donald N., and Nash, John (1984). "Corn at Interest: The Extent and Cost of Grain Storage in Medieval England." American Economic Review, 74(1): 174-187.

[52] Mitchell, B.R. and H.G. Jones (1971). "Second Abstract of British Historical Statistics." Cambridge: CUP.

[53] Mokyr, Joel (1977). "Demand vs. Supply in the Industrial Revolution." Journal of Economic History 37(4): 981-1008.

[54] Mokyr, Joel (1990). "The Lever of Riches: Technological Creativity and Economic Progress." New York: Oxford University Press.

[55] Mokyr, Joel (1993). "The New Economic History and the Industrial Revolution." In Joel Mokyr, ed., The British Industrial Revolution: an Economic Perspective. Boulder: Westview Press: 1-131.

[56] Mokyr, Joel (1999). "Editor's Introduction. The New Economic History and the Industrial Revolution." In: J. Mokyr, ed., The British Industrial Revolution: An Economic Perspective. Boulder, CO: Westview.

[57] Mokyr, Joel (2002). "Why was the Industrial Revolution a European Phenomenon?", Supreme Court Economic Review 9.

[58] Mokyr, Joel and Hans-Joachim Voth (2008). "Understanding Growth in Europe, 1700-1870 – First Industrializations, Unified Growth Theory, and Beyond." in: Steven Broadberry, Kevin O'Rourke, Cambridge Economic History of Europe, vol. I, Cambridge, forthcoming.

[59] Mosk, Carl (1976). "Fecundity, Infanticide, and Food Consumption in Japan." Explorations in Economic History 15(3): 269-89.

[60] Murphy, Kevin M, Andrei Shleifer and Robert W Vishny (1989a). "Income Distribution, Market Size, and Industrialization." Quarterly Journal of Economics 104(3): 537-64.

[61] Murphy, Kevin M., Andrei Shleifer and Robert Vishny (1989b). "Industrialization and the Big Push." Journal of Political Economy 97(5): 1003-26.

[62] North, Douglass C., und Robert P. Thomas (1973). "The Rise of the Western World: A New Economic History." Cambridge, Cambridge University Press.

[63] Pomeranz, Kenneth (2000). "The Great Divergence : China, Europe, and the Making of the Modern World Economy." Princeton, N.J., Princeton University Press.

[64] Romer, Paul M. (1990). "Endogenous Technological Change." Journal of Political Economy 98(5): 71-102.

[65] Rozman, G. (1973). "Urban Networks in Ch'ing China and Tokugawa Japan." Princeton, NJ: Princeton University Press.

[66] Stokey, Nancy (2001). "A Quantitative Model of the British Industrial Revolution, 1780-1850." Carnegie-Rochester Conference Series on Public Policy 55: 55-109.

[67] Strulik, H. (2006), "Patterns of Demographic Development and Structural Change", University of Copenhagen working paper.

[68] Temin, Peter and Hans-Joachim Voth (2005). "Credit Rationing and Crowding Out during the Industrial Revolution: Evidence from Hoare's Bank 1702–1862." Explorations in Economic History 42(3): 325-348.

[69] Voth, Hans-Joachim (2003). "Living Standards During the Industrial Revolution: An Economist's Guide." American Economic Review 93(2): 221-226.

[70] de Vries, Jan (1984). "European Urbanization 1500-1800." London: Methuen.

[71] Williamson, Jeffrey G. (1984). "Why was British growth so slow during the Industrial Revolution?" Journal of Economic History 44(3): 687–712.

[72] Wrigley, Edward Anthony (1983). "The Growth of Population in Eighteenth-Century England: A Conundrum Resolved." Past and Present 98: 121-50.

[73] Wrigley, Edward Anthony (1988). "Continuity, Chance and Change: The Character of the Industrial Revolution in England." Cambridge: Cambridge University Press.

[74] Wrigley, Edward Anthony, R.S. Davies, Jim Oeppen and Richard Schofield (1997). "English Population History from Family Reconstitution 1580-1837." Cambridge: CUP.

[75] Wrigley, Edward Anthony and Richard Schofield (1981). "The Population History of England, 1541-1871: A Reconstruction." Cambridge, MA, Harvard University Press.

[76] Zweimüller, Josef (2000). "Schumpeterian Entrepreneurs Meet Engel's Law: The Impact of Inequality on Innovation-Driven Growth." Journal of Economic Growth 5(2): 185-206.

Chapter 2

The Three Horsemen of Growth: Plague, War and Urbanization in Early Modern Europe

(Joint with Joachim Voth, UPF)

2.1 Introduction

In 1400, Europe's chances for rapid economic development seemed small. The continent was politically fragmented, torn by frequent military conflict, and dominated by feudal elites. Literacy was low. Other regions, such as China, appeared more promising. It had a track record of useful inventions, from ocean-going ships to gunpowder and advanced clocks (Moykr 1990). The country was politically unified, and governed by a career bureaucracy chosen by competitive exam (Pomeranz 2000). Few if any of the important variables analyzed in modern growth studies suggest that Europe looked promising.[1]

By 1700 however, and long before it industrialized, Europe had pulled ahead decisively - a first "Great Divergence" had occurred (Broadberry and Gupta 2006, Diamond 1997).[2] England's per capita income was more than twice that of China, European silver wages were often markedly higher, and European urbanization rates were more than double those in China (Broadberry and Gupta 2006, Maddison 2003). This early divergence matters in its own right. It laid the foundations for the European conquest of vast parts of the globe (Diamond 1997). More importantly, it may have contributed to the even greater differences in per capita incomes that followed. In many unified growth models, a gradual or temporary rise of per capita income is crucial for starting the transition to self-sustaining growth (Galor and Weil 2000, Hansen and Prescott 2002). Also, higher starting incomes may increase a country's industrialization probabilities (Voigtländer and Voth 2006). If we are to understand why Europe achieved the transition from "Malthus to Solow" before other regions of the world, it is necessary to explain this initial divergence of incomes.

In this paper, we identify a new puzzle, and argue that its solution can help explain why the most advanced parts of Europe were far ahead of the rest of the world by 1700 already. The early modern divergence in per capita incomes represents a major puzzle for Malthusian models because per capita incomes should not be able to rise substantially above subsistence for an extended period. Before industrialization, the 'fertility of wombs' was necessarily greater than the 'fertility of minds.' Galor (2006) estimates that TFP grew by no more than 0.05-0.15% p.a. in the pre-industrial era. Over a century, productivity could increase by 5-16%. Maximum fertility rates per female, by contrast, are around 7. Even with only 3 surviving children, a human population growing unconstrained would quadruple after 100 years.[3] This is why, in a Malthusian regime, past generations should have always, in HG Well's words,

[1] For a recent overview, see Bosworth and Collins (2003) and Sala-i-Martin et al. (2004).

[2] Pomeranz (2000), comparing the Yangtze Delta with England, argues the opposite. The consensus now is that his revisionist arguments to do no stand up to scrutiny (Allen 2004; Allen, Bengtsson, and Dribe 2005; Broadberry and Gupta 2006).

[3] Assuming a generation length of 25 years.

"spent the great gifts of science as rapidly as it got them in a mere insensate multiplication of the common life."[4]

Nonetheless, living standards in many European countries increased throughout the early modern period. Maddison (2007) estimates that Western European per capita incomes increased by more than 30%, and aggregate incomes still more between 1500 and 1700.[5] His figures are imperfect, but knowledgeable observers such as Adam Smith detected the same trend: "the annual produce of the land and labour of England... is certainly much greater than it was a little more than a century ago at the restoration of Charles II (1660)... and [it] was certainly much greater at the restoration than we can suppose it to have been a hundred years before."[6] How could such a marked rise be sustained over such a long period, despite the potential for rapid population growth to erode all gains quickly?

We argue that the impact of the Black Death in Europe was crucial. Western Europe's unique set of geographical and political starting conditions interacted with the plague shock to make higher per capita living standards sustainable. In a Malthusian regime, lower population spells higher wages. Because the shock was very large, with up to half of the population dying, land-labor ratios improved, and wages increased substantially. These real wage gains were so large, and concentrated in such a brief period of time, that they could not be undermined quickly by population growth. Wages remained high for more than one or two generations, and were partly spent on manufactured goods. Their production required a higher percentage of the labor force in the urban sector. Because early modern European cities were death-traps with mortality far exceeding fertility rates, they would have disappeared had it not been for steady in-migration from the countryside. Thus, the extra demand for manufactures pushed up average death rates, making higher incomes sustainable. We capture these key elements in a simple two-sector model. Effectively, Engel's law ensured that the plague's positive effect on wages did not wear off entirely as a result of higher fertility and lower mortality. Because changes in the composition of demand increased urbanization rates, average death rates became permanently higher, making the wage gains sustainable.

This benign effect was reinforced because city wealth fueled early modern Europe's endemic warfare. Between 1500 and 1800, the continent's great powers were fighting each other on average for nine years out of every ten (Tilly 1990). Cities also acted as nuclei for long-distance trading networks. Both war and trade spread epidemics. The more effectively they did so, the higher death rates overall were, and the more readily a rise in incomes and in the urban share of the population could be sustained. In this way, three "Horsemen of Death" - plague, war, and urbanization - led to higher incomes. The combination of these three factors is what we call the European Mortality Pattern. In contrast to numerous papers identifying a negative (short-run) effect of wars, civil wars, disease, and epidemics on growth in economies today,[7] we argue that they acted as "Horsemen of Growth".

The great 13$^{\text{th}}$ century plague also affected China, as well as other parts of the world (McNeill 1977). Why did it not have the same effects? We argue that two factors were crucial. Chinese cities were far healthier than European ones, for a number of reasons involving cultural practices and political con-

[4] Wells 1905. Galor and Weil (2000) assume that the response of fertility to incomes is delayed. Hence, a one-period acceleration in technological change can generate higher incomes in the subsequent period, and a sequence of positive shocks can lead to sustained growth. While this solves the problem in a technical sense, it is unlikely to explain why fertility responses did not erode real wage gains over hundreds of years.

[5] Maddison estimates that total real GDP doubled in the same period.

[6] Smith, 1776 (1976), pp. 365-66.

[7] Murdoch and Sandler 2002, Hoeffler and Querol-Reynal 2003, Hess 2003.

ditions. Also, political fragmentation in Europe ensured that greater wealth in cities helped to finance almost continuous warfare after 1500. Since China was politically unified, there was no link between city growth and the frequency of armed conflict. Hence, a very similar shock did not lead to permanently higher death rates; per capita incomes could not rise.[8]

The mechanism presented in this paper is not the only one that can deliver a divergence in per capita incomes without technological change. In addition to high death rates, Europeans curtailed birth rates. In contrast to many other regions of the world, socio-economic factors, and not biological fertility, determined the age at first marriage for women. This is what Hajnal (1965) termed the European Marriage Pattern. In our calibrations, we find that fertility restriction can explain part of the European advantage, but that the mortality effects identified in our model account for more than half of the "Great Divergence".

We are not the first to argue that higher death rates can have beneficial economic effects. Young (2005) concludes that Aids in Africa has a silver lining because it reduces fertility rates, increasing the scarcity of labor and thereby boosting future consumption. Lagerlöf (2003) also examines the interplay of growth and epidemics, but argues for the opposite causal mechanism. He concludes that a decline in the severity of epidemics can stimulate growth if they stimulate population growth and human capital acquisition. Brainard and Siegler (2003) study the outbreak of "Spanish flu" in the US, and conclude that the states worst-hit in 1918 grew markedly faster subsequently. Compared to these papers, we make two contributions. We are the first to construct a consistent model demonstrating how specific European characteristics - political and geographical - interacted with a mortality shock to drive up living standards over the long run. Also, we calibrate our model to show that it can account for a large part of the "Great Divergence" in the early modern period.

Other related literature includes the unified growth models of Galor and Weil (2000), and Galor and Moav (2003). In both, before fertility limitation sets in and growth becomes rapid, a state variable gradually changes over time during the Malthusian regime, making the final escape from stagnation more and more likely. In Galor and Weil (2000) and in Jones (2001), the rise in population which in turn produces more ideas is a key factor; in Galor and Moav (2003), it is the quality of the population.[9] Hansen and Prescott (2002) assume that productivity in the manufacturing sector increases exogenously, until part of the workforce switches out of agriculture. Our model abstracts from technological change during the Malthusian era, and emphasizes changes in death rates as a key determinant of living standards. One of the key advantages is that it can be applied to the cross-section of growth. In contrast, the majority of unified growth papers implicitly uses the world as their unit of observation.

We proceed as follows. The next section provides a detailed discussion of the historical context. Section 3 introduces a simple two-sector model that highlights the main mechanisms. In Section 4, we calibrate our model and show that it captures the salient features of the "Great Divergence", compare the effect of the European Mortality Pattern to the consequences of fertility restriction, and compare the model predictions with actual data. The final section summarizes our findings and puts them in the context of explanations of the transition to self-sustaining growth.

[8]Hui (2005) compares the Warring States period in China (656-221 BC) with early modern Europe, and argues that flawed strategy is largely to blame for Europe's failure to unify politically.

[9]Clark (2007) finds some evidence in favor of the Galor-Moav hypothesis, with the rich having more surviving offspring.

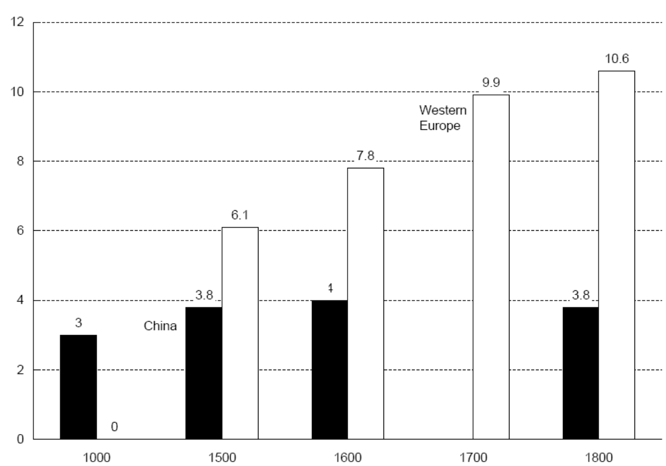
2.2 Historical context and background

Our story emphasizes three elements that can explain the first "Great Divergence": the impact of the plague, the peculiarities of European cities, and interaction effects with the political environment. In this section, we first assemble some of the evidence suggesting that European growth during the early modern period was unusually rapid, and then discuss the three central elements in our model in turn.

The Great Divergence

That Europe pulled ahead of the rest of the world in terms of per capita living standards is now a widely accepted fact. While Pomeranz (2000) argued that farmers in the Yangtze delta in China earned the same wage in terms of calories as English farmers, there is now a broad consensus that overturns this argument. First, better data strongly suggest that English wages expressed as units of grain or rice were markedly higher. Broadberry and Gupta (2006) calculate Chinese grain-equivalent wages were 87% of English ones in 1550-1649, and fell to 38% in 1750-1849. Second, since foodstuffs were largely non-traded goods, they are a poor basis for comparison. Silver wages were much higher in Europe than in China. According to Broadberry and Gupta, they fell from 39% of the English wage to a mere 15%.[10] Finally, urbanization rates have been widely used as an indicator of economic development (Acemoglu, Johnson and Robinson 2005). They strongly suggest that Europe overtook China at some point between 1300 and 1500, and then continued to extend its lead (figure 3.1).

Figure 2.1: Urbanization rates in China and Europe, 1000-1800. Source: Maddison 2003



The beneficial effect of the Black Death on real wages is well-documented. The wage figures for England by Phelps-Brown and by Clark (2005) suggest that wages broadly doubled after 1350. If and when these gains were reversed, and to what extent, is less clear. The older Phelps-Brown series suggests a strong reversal. Clark (2005) shows that wages fell back from their peak somewhat, but except for crisis

[10]While Broadberry and Gupta's figures for the second period are partly influenced by values from the early 19th century, when industrialization was already under way, it is clear that observations for the 18th century alone would also show a marked advantage.

years around the English Civil War, they remained about fifty percent above their 1300 level.[11] In this sense, they offer some indirect support to the optimistic GDP figures provided by Maddison (2003).

Changes in Europe were not uniform. Allen (2001) found that the real wage gains for craftsmen after the Black Death were only maintained in Northwestern Europe. In Southern Europe - especially Italy, but also Spain - stagnation and decline after 1500 are more noticeable. Yet for every single European country with the exception of Italy, Maddison estimates that per capita GDP was higher by 1700 than it had been in 1500. This indirectly suggests that standard Malthusian predictions did not hold during the period. Maddison argues that subsistence is equivalent to ca. $400 US-Geary Khamy dollars. Even relatively poor countries like Spain and Portugal had per capita incomes more than twice as high in 1700. At this stage, every single European country had been above the threshold for centuries, often by 50 percent or more. This is the puzzle that we seek to explain.

The Plague

The plague arrived in Europe from the Crimea in December 1347. Besieging Tartar troops suffered from the disease. In an early example of biological warfare, they used catapults to throw bodies of the deceased over the city wall of Caffa, a Genoese trading outpost. Soon, the city population caught the disease. It spread via the shipping routes, first to Constantinople, then to Sicily and Marseille, then mainland Italy, and finally the rest of Europe. By December 1350, it had spread to the North of England and the Baltic (McNeill 1977).

Mortality rates amongst those infected varied from 30 to 95%. Bubonic and pneumonic forms of the plague both contributed to surging mortality. The bubonic form was transmitted by fleas and rats carrying the plague bacterium (Yersinia pestis). Infected fleas would spread the disease from one host to the next. When rats died, fleas tried to feast on humans, infecting them in the process. In contrast, pneumonic plague spread person to person, by coughing of the infected. Transmission and mortality rates were particularly high for this form of the plague.

There appear to have been few differences in mortality rates between social classes, or between rural and urban areas. Some city-dwellers tried to escape the plague, by withdrawing to country residences, as described in Decamerone. It is unclear how often these efforts succeeded. Only a handful of areas in the Low Countries, in Southwest France and in Eastern Europe were spared the effects of the Black Death.

We do not have good estimates of aggregate mortality for medieval Europe. Most estimates put population losses at 15 - 25 mio., out of a total population of approximately 40 mio. people. Approximately half of the English clergy died, and in Florence and Venice, death rates have been estimated as high as 60-75%.

City Mortality

European cities were deadly places. In 1841, when large inflows of labor put particular pressure on urban infrastructures, life expectancy in Manchester was a mere 25 years. At the same time, the national average was 42, and in rural Surrey, 45 years. Early modern cities were often equally unhealthy. Life expectancy

[11] What matters for the predictions of the Malthusian model is per capita output, not wages as such. National income in the aggregate will be equivalent to the sum of wages, rents, and capital payments. Since English population surpassed its 1300 level in the eighteenth century, it is likely that rental payments were higher, too.

in London, 1580-1799, fluctuated between 27 and 28 years (Landers 1993). Nor were provincial towns much more fortunate. York had similar rates of infant mortality.[12] In France, the practice of wet-nursing (sending children from cities for breast-feeding to the countryside) complicates comparisons. A comprehensive survey of rural-urban mortality differences estimates that in early modern Europe, life expectancy was 1.5 times higher in the countryside (Woods 2003).

Mortality figures for China have been reconstructed based on the family trees of clans (Tsui-Jung 1990). Infant mortality rates were lower in cities than in rural areas, and life expectancy was higher. While the data is not necessarily representative, other evidence lends indirect support. For example, life expectancy in Beijing in the 1920s and 1930s was higher than in the countryside. Members of Beijing's elite in the 18th century experienced infant mortality rates that were half those in France or England (Woods 2003). Given that, in Europe at least, class differences in mortality were not common in cities, there is a good chance that mortality rates in general might have been low.

In Japan, where some data for 18th century Nakahara and some rural villages survives, city dwellers lived as long as their cousins in the countryside. Some recent evidence (Hayami 2001) on adult mortality questions if cities were indeed healthier than the countryside, as some scholars have argued (Hanley 1997; Macfarlane 1997). What is clear is that on balance, the evidence favors the hypothesis that there was no large urban penalty in the Far East. The main reasons probably include the transfer of "night soil" (i.e., human excrement) out of the city and onto the surrounding fields for fertilization, high standards of personal hygiene, and a diet that emphasized vegetarian food. Since the proximity of animals is a major cause of disease, all these factors probably combined to reduce the urban mortality burden.

In the view of one prominent urban historian, in "1600, just as in 1300, Europe was full of cities girded by walls and moats, bristling with the towers of churches." (DeVries 1976). In China, city walls were widely used throughout the early modern period, partly because of their symbolic value for administrative centers of the Empire. However, since the country was unified under the Qin Dynasty, the defensive function of city walls declined. With relative ease, houses and markets spread outside the city walls.[13] Because Far Eastern cities could easily expand beyond the old fortifications, city growth did not push up population densities in the same way as in Europe.[14]

In many European countries, regulations further ensured that manufacturing activities and market exchange was largely a monopoly of the cities.[15] In China, periodic markets in the countryside served the same function, reducing relative urbanization rates (Rozman 1973). Finally, European cities offered a unique benefit not found in other parts of the world - the chance to escape servitude. The general rule of staying within the city walls for one year and one day made free men out of peasants bound to the land and their lord. In contrast, "Chinese air made nobody free".[16]

[12] Galley 1998. There is not enough data to derive life expectancy. Since infant mortality is a prime determinant, it was probably in the same range.

[13] In some cases, the new suburbs would also be enclosed by city walls (Chang 1970).

[14] Barcelona is one extreme example. After the 1713 uprising, the Bourbon kings did not allow the city to expand beyond its existing walls until 1854. As industrial growth led to an inflow of migrants, living conditions deteriorated considerably (Hughes 1992).

[15] Some scholars have argued that "proto-industrialization", i.e. early forms of home-based manufacturing, often located in the countryside, were an important feature of early modern European growth (Ogilvy and Cerman 1996). This view is not widely accepted (Coleman 1983).

[16] Mark Elvin, cited in Bairoch (1991).

Wars, Trade and Disease

The available data on deaths caused by military operations in the early modern period is sketchy.[17] What is clear is that diseases spread by armies were far more important than battlefield casualties and the deaths of siege victims in determining mortality rates. While individual campaigns could be deadly, armies were too small, and their members too old, to influence aggregate mortality rates significantly.[18] The plague of 1347-48 was not the last to strike Europe. In the period 1347-1536, there were outbreaks every 7 years. Until the 1670s, frequency declined by half. The last incidents in Western Europe were plague outbreaks in Austria (1710) and Marseille (1720). Warfare and the outbreak of diseases were closely linked. The Black Death had originally arrived with a besieging Tartar army in the Crimea. Early modern armies killed many more Europeans by the germs they spread than through warfare. Isolated communities in the countryside would suddenly be exposed to new germs as soldiers foraged or were billeted in farmhouses. The effect could be as deadly as it had been in the New World, where European diseases killed millions. In one famous example, it has been estimated that a single army of 6,000 men, dispatched from La Rochelle to deal with the Mantuan Succession, spread plague that killed over a million people (Landers 2003). Population losses in the aggregate could be heavy. The Holy Roman Empire lost 5-6 mio. out of 15 mio. inhabitants during the Thirty Years War; France lost 20% of its population in the late 16th century as a result of civil war. As late as in the Napoleonic wars, typhus, smallpox and other diseases spread by armies marauding across Europe proved far deadlier than guns and swords.

For the early and mid-nineteenth century, we have data that allows some gauging of the orders of magnitude involved. In the Swedish-Russian war of 1808-09, mortality rates in Sweden doubled, almost exclusively through disease. In isolated islands, the presence of Russian troops led to a tripling of death rates. During the Franco-Prussian and the Austro-Prussian wars later in the 19th century, non-violent death rates increased countrywide by 40-50% (Landes 2003). Both background mortality and the impact of war were probably lower than in the early modern period. Warfare was less likely to spread new germs, since in areas touched by troop movements were now integrated by extensive railway networks. The figures for the Thirty Years War and for 16th century France similarly suggest increases in mortality above their normal rate by 50 to 100%.

Early modern warfare, with its need for professional, drilled troops, Italian-style fortifications, ships, muskets and cannons were particularly expensive – money formed the sinews of power (Brewer 1990, Landers 2003). To fight wars, princes needed access to liquid wealth. Philip II's silver allowed him to fight a war in every year of his reign except one. Elsewhere, the growth of cities provided the kind of easily mobilized wealth that could be spent on mercenary armies - either directly, through taxation, or through sovereign lending. With the growth of urbanization in early modern Europe, the financial means for fighting more, fighting longer, and in more deadly fashion became more easily accessible.

China in the early modern period saw markedly less warfare than Europe. We calculate that even on the most generous definition, wars and armed uprisings only occurred in one year out of five, no more than a quarter of the European frequency. Not only were wars fewer in number. They also produced less of a

[17] Landers (2003) offers an overview of battle-field deaths.

[18] Since infant mortality was high, by the time men could join the army, many male children had died already. This makes it less likely for military deaths to matter in the aggregate. Lindegren (2000) finds that military deaths only raised Sweden's death rates by 2-3/1000 in most decades between 1620 and 1719, a rise of no more than 5%. Castilian military deaths were 1.3/1000, equivalent to 10 percent of adult male deaths but no more than 3-4% of overall deaths.

spike in epidemics. Europe is geographically subdivided by rugged mountain ranges and large rivers, with considerable variation in climatic conditions. China overall is more homogenous in geographical terms. While rugged in many parts, major population centers were not separated by geographical barriers in the same way as in Europe. Since linking semi-independent disease pools through migratory movements pushes up death rates in a particularly effective way, it may also be that in every armed conflict, similar troop movements produced less of a surge in Chinese death rates than in Europe.[19]

Compared to warfare, trade in early modern Europe was a less effective, but more frequent cause of disease spreading. This is why quarantine measures became frequent throughout the continent. The last outbreak of the plague in Europe occurred in Marseille in 1720. A plague ship from the Levant, with numerous sufferers on board, was first quarantined, only to have the restriction lifted as a result of commercial pressure. It is estimated that 50,000 out of 90,000 inhabitants died in the subsequent outbreak (Mullett 1936). Since trade increases with per capita incomes, the positive, indirect effect of the initial plague on wages created knock-on effects. These combined to raise mortality rates yet further. In addition, there were interaction effects between the channels we have highlighted. The effectiveness of quarantine controls, for example, often declined when wars disrupted administrative procedure (Slack 1981).

2.3 The Model

This section presents a simple two-sector model that captures the basic mechanisms determining pre-industrial living standards. The economy is composed of N identical individuals who work, consume, and procreate. N_A individuals work in agriculture (A) and live in the countryside, while N_M agents live in cities producing manufacturing output (M), both under perfect competition.[20] For simplicity, we assume that wages are the only source of income. Agents choose their workplace in order to maximize expected utility, trading greater risks of death in the city for a higher wage. Agricultural output is produced using labor and a fixed land area. This implies decreasing returns in food production. Manufacturing uses labor only and is subject to constant returns to scale. Preferences over the two goods are non-homothetic and reflect Engel's law: The share of manufacturing expenditures (and thus the urbanization rate N_M/N) grows with income.

Population growth responds to per-capita income. Higher wages translate into more births and lower mortality. Therefore, the economy is Malthusian – per capita income stagnates close to the subsistence level, keeping most people at the edge of starvation ("positive" Malthusian check). With stagnating technology, death rates equal birth rates, and N is constant in equilibrium. Technological progress temporarily relieves Malthusian constraints; population can grow. In the absence of ongoing productivity gains, however, the falling land-labor ratio drives wages back to their original equilibrium level. Per-capita income is therefore self-equilibrating.

An epidemic like the plague has an economic effect akin to technological progress: it causes land-labor ratios to rise dramatically. This leaves the remaining population with greater per-capita income,

[19] We are indebted to David Weil for this point. Weil (2004) shows the marked similarity of agricultural conditions in large parts of modern-day China.

[20] During the early modern period, a substantial share of manufacturing took place outside cities – a process called "protoindustrialization" by some. We abstract from it since cities still grew, and our key mechanism remains intact, even if some of the additional demand translated into growth for non-urban manufactured goods.

which translates into more demand for manufactured goods. As a consequence, urbanization rates have to rise. In the absence of productivity growth and shifts in the birth or death schedules, subsequent population growth pulls the economy back to its earlier equilibrium – there is no escape from Malthusian stagnation. However, after the plague, the 'Horsemen of Death' start to ride high: Wars become more frequent. City mortality is high. Increasing trade, linking the urban nuclei, spreads disease, as do wars. As these become a permanent feature of the early modern European economy, the death schedule shifts upwards. We argue that this mechanism captures an important element of the European experience in the centuries between the Black Death and the Industrial Revolution. The new long-run equilibrium has higher birth and death rates, but also increased per capita incomes and a higher share of the population living in cities.

2.3.1 Consumption

Each individual supplies one unit of labor inelastically in every period. There is no investment – individuals i use all their income to consume homogenous agricultural goods ($c_{A,i}$) and manufactured goods ($c_{M,i}$). At the beginning of each period, agents choose their workplace in order to maximize expected utility. Agents' optimization therefore involves two stages: The choice of their workplace and the optimal spending of the corresponding income. We consider the latter first.

In the intra-temporal optimization, each individual takes workplace-specific wages $w_i,\ i = \{A, M\}$ as given and maximizes instantaneous utility.[21] The corresponding budget constraint is $c_{A,i} + p_M c_{M,i} \leq w_i$, where p_M is the price of the manufactured good. The agricultural good serves as the numeraire. Before they begin to demand manufactured goods, individuals need to consume a minimum quantity of food, $\underline{c}$. In the following, we refer to this number as the subsistence level, meaning that individuals satisfy their basic needs for calories at $\underline{c}$. Below $\underline{c}$, individuals suffer from hunger, but do not necessarily die – mortality increases continuously as c_A falls. Preferences take the Stone-Geary form and imply the composite consumption index:

$$u(c_A, c_M) = \begin{cases} (c_A - \underline{c})^{\alpha} c_M^{1-\alpha}, & \text{if } w_i > \underline{c} \\ \beta(c_A - \underline{c}), & \text{if } w_i \leq \underline{c} \end{cases} \tag{2.1}$$

Where $\beta > 0$ is a parameter specified below. Given w_i, consumers maximize (2.1) subject to their budget constraint. In a poor economy, where income is not enough to ensure subsistence consumption $\underline{c}$, equation (2.2) does not apply. In this case, the starving peasants are unwilling to trade food for manufactured goods such that the relative price p_M and rural wages w_M are zero. Thus, there are no cities and all individuals work in the countryside: $N_A = N$, while $c_A = w_A < \underline{c}$.

When agricultural productivity is large enough to provide above-subsistence consumption $w_A > \underline{c}$, expenditure shares on agricultural and manufacturing products are:

$$\begin{aligned} \frac{c_{A,i}}{w_i} &= \alpha + (1-\alpha)\left(\frac{\underline{c}}{w_i}\right) \\ \frac{p_M c_{M,i}}{w_i} &= (1-\alpha) - (1-\alpha)\left(\frac{\underline{c}}{w_i}\right) \end{aligned} \tag{2.2}$$

[21] In the following, the subscripts A and M not only represent agricultural and manufacturing goods, but also the locations of production, i.e., countryside and cities, respectively.

Once consumption passes the subsistence level, peasants start to demand manufacturing products, which leads to the formation of cities. If income grows further, the share of spending on manufactured goods grows in line with Engel's law, and cities grow in size. The relationship between income and urbanization is governed by the parameter α. A larger α implies more food expenditures and thus less urbanization at any given income level.

2.3.2 Production

Agricultural and manufactured goods are homogenous, and are produced under constant returns and perfect competition. In the countryside, peasants use labor N_A and land L to produce food. The agricultural production function is

$$Y_A = A_A N_A^{\gamma} L^{1-\gamma} \tag{2.3}$$

where A_A is a productivity parameter and γ is the labor income share in agriculture. Suppose that there are no property rights over land. Thus, the return to land is zero and agricultural wages are equal to the average product of labor:

$$w_A = A_A \left(\frac{L}{N_A}\right)^{1-\gamma} = A_A \left(\frac{l}{n_A}\right)^{1-\gamma} \tag{2.4}$$

where $l = L/N$ is the land-labor ratio and $n_A = N_A/N$ is the labor share in agriculture, or rural population share. Since land supply is fixed, increases in population result in a falling land-labor ratio and ceteris paribus in declining agricultural wages. Manufacturing goods are produced in cities using the technology

$$Y_M = A_M N_M \tag{2.5}$$

where A_M is a productivity parameter. Manufacturing firms maximize profits and pay wages $w_M = p_M A_M$. The manufacturing labor share n_M is identical to the urban population share.

2.3.3 Migration

The optimal workplace choice determines migration in our model. We suppose that migration occurs at the beginning of every period t, such that migrating individuals arrive at their workplace before production starts in t. In early modern Europe, death rates were substantially higher in cities as compared to rural areas ($d_M > d_A$). Higher mortality rates in the cities were compensated by higher average wages. In order to set up the corresponding optimization problem, we first derive the indirect utility of consumers from (2.1) and (2.2):

$$\widetilde{u}(w_i, p_M) = \left(\frac{1}{p_M}\right)^{1-\alpha} \alpha^{\alpha} (1-\alpha)^{1-\alpha} (w_i - \underline{c}) \tag{2.6}$$

Note that this equation is valid only if ($w_i > \underline{c}$), which is the more interesting case on which we concentrate from now on. Individuals maximize expected utility in each period, where $(1 - d_i)$ is the survival probability when working at place $i = \{A, M\}$. We define the (hypothetical) utility associated with death as the one corresponding to zero consumption: $\widetilde{u}(0, p_M) = -\beta\underline{c}$, as implied by (2.1).[22] For the

[22] Any negative number associated with the utility level of death serves to obtain a positive city wage premium – the more negative, the higher the premium.

following steps it is convenient to define $\beta \equiv (1/p_M)^{1-\alpha}\, \alpha^\alpha\, (1-\alpha)^{1-\alpha}$. The optimization problem is then:

$$\max_{i=\{A,M\}} \{(1-d_i)\, \widetilde{u}(w_i, p_M) \; + d_i\, \widetilde{u}(0, p_M)\} \tag{2.7}$$

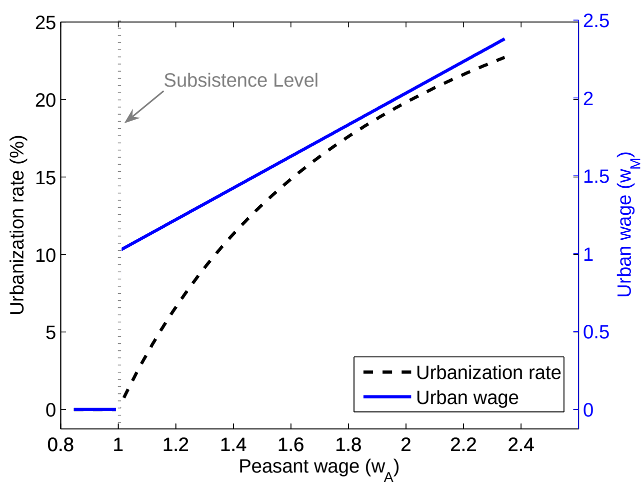
This setup implies that the city and countryside expected utility levels are equal whenever no migration is desired. In this case, (2.7) yields:

$$(w_M - \underline{c}) = \frac{(1-d_A)}{(1-d_M)}(w_A - \underline{c}) + \frac{d_M - d_A}{1-d_M}\, \underline{c} \tag{2.8}$$

Since $d_M > d_A$, wages in the city are higher than in the countryside.[23] If (2.8) holds with equality, no migration occurs. When the LHS is larger than the RHS, the urban wage premium outweighs the excess mortality in cities, attracting rural workers. The rising urban labor supply then causes the relative wage to drop until equality is re-established. The opposite workplace decisions restore the equilibrium when the RHS is larger than the LHS. These dynamics can immediately correct minor shocks to relative productivity or population N_A and N_M. If shocks are large, like the plague, migration must be large to re-establish equality in (2.8). In this case, cities grow less than would be predicted by the baseline model as it takes time to build urban infrastructure – Rome was not built in a day. We discuss this case in detail in section 2.3.5.

Figure 3.2 illustrates the basic income-demand-urbanization mechanism of our model. If the rural wage (horizontal axis) is below subsistence, the starving population does not demand any manufacturing goods and cities do not exist (zero urbanization, left axis). Correspondingly, there are no manufacturing workers (zero urban wages, right axis). Cities emerge once peasants' productivity is large enough to provide above-subsistence consumption, such that agents also demand manufacturing goods. At the same time, urban consumption becomes important, driven by city workers who produce manufacturing output. As productivity grows further, urbanization and consumption (both urban and rural) grow in tandem.

Figure 2.2: Wages and urbanization



[23] If rural income is too small to ensure consumption above subsistence ($w_A \leq \underline{c}$), equation (2.8) does not hold and there is no migration, since all agent work in agriculture.

2.3.4 Population Dynamics

Birth and death rates depend on real p.c. income. Since there is no investment, units of consumption serve as a measure of real income: $c_{\bullet,i} = c_{A,i} + c_{M,i}$ for $i = \{A, M\}$.[24] Substituting from (2.2) into this expression yields:

$$c_{\bullet,i} = \alpha w_i + (1-\alpha)\underline{c} + \frac{(1-\alpha)}{p_M}(w_i - \underline{c}) \tag{2.9}$$

Individuals at location i procreate at the birth rate

$$b_i = b_0 \cdot (c_{\bullet,i})^{\varphi_b} \tag{2.10}$$

where $\varphi_b > 0$ is the elasticity of the birth rate with respect to real income. Note that $c_{\bullet,i} = \underline{c}$ if $w_i = \underline{c}$. We choose $\underline{c} = 1$, so that b_0 represents the birth rate at subsistence income. Before the Black Death, location-specific death rates fall with income and are given by

$$\begin{aligned} d_A &= \min\{1,\ d_0 \cdot (c_{\bullet,A})^{\varphi_d}\} \\ d_M &= \min\{1,\ d_0 \cdot (c_{\bullet,M})^{\varphi_d} + \triangle d_M\} \end{aligned} \tag{2.11}$$

where $\varphi_d < 0$ is the elasticity of the death rate with respect to real income and $\triangle d_M$ represents city excess mortality; d_0 is the countryside death rate at subsistence income.

Higher p.c. income and urbanization after the plague spur trade and wars. Military casualties mount. Armies as well as merchants continuously spread pathogenic germs across cities and countryside. These factors raise background mortality. In combination, this is what we call the 'Horsemen effect', h. Because it is driven by growing income and urbanization, we use the urbanization rate n_M as a proxy for its strength. To capture the positive relationship between urbanization and Horsemen mortality, we calculate h as:

$$h(n_M) = \begin{cases} 0, & \text{if } n_M \leq n_M^h \\ \min\{\delta\, n_M,\ h_{max}\}, & \text{if } n_M > n_M^h \end{cases} \tag{2.12}$$

where $\delta > 0$ is a slope parameter, h_{max} represents the maximum additional mortality due to the Horsemen effect, and n_M^h is the threshold urbanization rate where the effect sets in. A poor economy with little urbanization has neither long-range mobility due to trade nor means for warfare; germ pools remain isolated and mortality is only driven by individual rural income as given by (2.11).[25] The role of the plague in our model is to introduce germs and to push p.c. income to levels where $n_M > n_M^h$. Neither germs nor higher income (and thus mobility) alone have an effect on long-run income levels. Only if higher mobility spreads epidemics, background mortality increases and alleviates the population pressure.

The last step before analyzing equilibria is to derive population growth from economy-wide fertility and mortality rates. We derive average fertility from (2.10), using the workforce shares n_A and n_M as weights:

$$b = n_A b_A + n_M b_M \tag{2.13}$$

[24] A simplified approach would have birth and death rate as functions of nominal income w_i, not taking into account changes in the relative price p_M. Because the latter changes substantially with the land-labor ratio, we choose the real income approach.

[25] A more detailed justification for $n_M^h > 0$ is that it indicates a minimum income level that cannot be expropriated, containing food for elementary nutrition as well as basic cloth and tools produced in city manufacturing. Once this threshold is passed, taxation yields the means for warfare and arouses the Horsemen.

The same method yields average death rates from (2.11) and (2.12), depending on whether or not the Horsemen are at work.

$$d = \begin{cases} n_A d_A + n_M d_M, & \text{if } n_M \leq n_M^h \\ n_A d_A + n_M d_M + h, & \text{if } n_M > n_M^h \end{cases} \tag{2.14}$$

Note that increasing real income has an ambiguous effect on mortality: Larger $c_{\bullet,i}$ translates into smaller death rates in (2.11). On the other hand, manufacturing demand rises with income, driving more people into cities where mortality is higher. Moreover, in the presence of the Horsemen effect, urbanization (proxying for the spread of epidemics through trade and wars) also implies larger overall background mortality. The aggregate impact of productivity on mortality depends on the model parameters (as shown in section 2.4.1).

Population growth equals the difference between the average birth and death rate: $\gamma_{N,t} = b_t - d_t$. The law of motion for aggregate population N is thus

$$N_{t+1} = (1 + b_t - d_t)N_t \tag{2.15}$$

Births and deaths occur at the end of a period, such that all individuals N_t enter the workforce in period t.

2.3.5 Equilibria

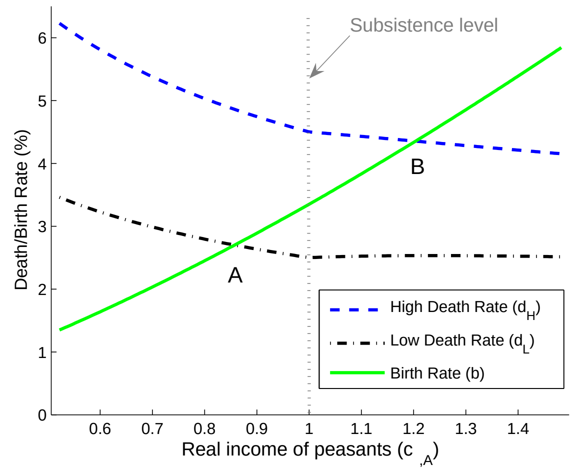
Equilibrium in our model is a sequence of factor prices, goods prices, and quantities that satisfies the intra-temporal and workplace optimization problems for consumers and firms. In this section, we analyze the economy without technological progress. The long-run equilibrium is characterized by stagnant population, labor shares, wages, prices, and consumption. All depend on how the birth and death schedule respond to income. Figure 3.3 visualizes this relationship. Real peasants' income $c_{\bullet,A}$ is shown on the horizontal axis.[26] Relatively low death rates lead to equilibrium A: a poor economy with below-subsistence income ($c_{\bullet,A} \leq \underline{c}$) where all individuals work in agriculture. The long-run level of consumption is independent of productivity parameters; it only depends on the intersection of b and d_L. For purposes of illustration, assume that there is a one-time major innovation in agriculture, augmenting A_A in equation (2.4). The rising wage shifts $c_{\bullet,A}$ to the right of point A, such that population grows ($b > d_L$). Consequently, the land-labor ratio l declines. So do wages, which eventually drives the economy back to equilibrium A. Land per worker is therefore endogenously determined in the long-run equilibrium.

In the absence of ongoing technological progress, there are two ways for achieving a permanent rise in per-capita income.[27] First, a permanent drop in birth rates, for example due to the emergence of the European Marriage Pattern. And second, a permanent increase in mortality – the main focus of this paper. Permanently higher death rates (d_H) imply lower population in equilibrium and therefore higher income,

[26]Provided that there is demand for manufacturing products, urban income is proportional to its rural counterpart, as shown in Figure 3.2.

[27]Continuous technological progress constantly pushes consumption to the right of point A, with increasing population and falling l always pulling it back. The equilibrium is thus located to the right of the intersection of b and d and is characterized by consumption stagnating at a higher level. Consumption can grow continuously only if technological change outpaces the falling land-labor ratio – a highly unrealistic scenario given the observed productivity growth of about 0.1% p.a. before the Industrial Revolution.

Figure 2.3: Population dynamics and equilibria



as represented by point B in Figure 3.3.[28] While total population is constant in point B, there must be perpetual migration from the countryside to cities in order to compensate city excess mortality.

Points A and B in Figure 3.3 are long-run equilibria with endogenous population size. For given technology A_A, productivity is fixed in the long-run, given by the endogenously determined land-labor ratio. During the transition to long-run equilibria, population dynamics change land per worker and thus productivity. In the following, we analyze these short-run equilibria. We first concentrate on the economy with below-subsistence consumption where individuals struggle for survival and produce only food in the countryside. Next, we turn to the economy with consumption above $\underline{c}$, accounting for constraints to migration due to city congestion during the transition process.

The Economy with Below-Subsistence Consumption

In order to check whether overall productivity (determined by A_A and the land-labor ratio) is sufficient to ensure above-subsistence consumption, we construct the indicator $\widehat{w}$, supposing that all individuals work in agriculture. Equation (2.3) then gives the corresponding per-capita income:

$$\widehat{w} \equiv \frac{Y_A(N)}{N} = A_A \left(\frac{L}{N}\right)^{1-\gamma} \tag{2.16}$$

If $\widehat{w} \leq \underline{c}$, all individuals work in agriculture ($N_A = N$) and spend their complete income on food. Since there is no demand for manufacturing goods, the manufacturing price is zero, implying zero urban wages and population. Economy-average fertility and mortality are thus equal to the rural levels given by equations (2.10) and (2.11).[29] Finally, there is no migration. In order to derive the long-run equilibrium, we calculate birth and death rates according to the equations in section 2.3.4. The intersection of the

[28] Note that the death schedules d_L and d_H become flatter when consumption passes the subsistence level. This is because richer agents also demand manufacturing products such that part of the population lives in cities, where mortality is higher ($\triangle d_M > 0$).

[29] Note that the Horsemen effect is zero because $n_M = 0$.

two schedules (point A in Figure 3.3) determines equilibrium income, which we can use to derive the corresponding population size N from (2.16).

Above-Subsistence Consumption and Unconstrained Migration

If $\widehat{w} > \underline{c}$, agricultural productivity is large enough to provide above-subsistence consumption. Following (2.2), the well-nourished individuals spend part of their income for manufacturing goods. Thus, a share n_M of the population lives and works in cities. In each period, individuals choose their profession and workplace based on their observation of income and mortality in cities and the countryside. Productivity increases lead to more manufacturing demand and spur migration, which occurs until (2.8) holds with equality. For small productivity changes, migration is minor and cities can absorb sufficiently many migrants to establish this equality immediately. We refer to this case as equilibrium with unconstrained migration. Goods market clearing together with equations (2.2), (2.3), and (2.5) implies

$$A_A N_A^{\gamma} L^{1-\gamma} = \alpha\left[(w_A - \underline{c})N_A + (w_M - \underline{c})N_M\right] + \underline{c}N \tag{2.17}$$

$$p_M A_M N_M = (1-\alpha)\left[(w_A - \underline{c})N_A + (w_M - \underline{c})N_M\right] \tag{2.18}$$

Solving for the expression in brackets in (2.18), plugging it into (2.17), and substituting $w_M = p_M A_M$ yields

$$\alpha w_M(1-n_A) + (1-\alpha)\underline{c} = (1-\alpha)A_A n_A^{\gamma} l^{1-\gamma} \tag{E1}$$

This equation contains two unknowns: n_A and w_M. We find an expression for the latter by using the equality in (2.8), as implied by unconstrained migration. Substituting (2.4) into (2.8) and rearranging gives:

$$(w_M - \underline{c}) = \frac{(1-d_A)}{(1-d_M)}\left[A_A\left(\frac{l}{n_A}\right)^{1-\gamma} - \underline{c}\right] + \frac{d_M - d_A}{1-d_M}\,\underline{c} \tag{E2}$$

We now need d_A and d_M as functions of n_A and w_M. Plugging (2.9) into (2.11), with w_A substituted from (2.4) and $p_M = w_M/A_M$, we obtain:

$$d_A = d_0\left[\alpha A_A\left(\frac{l}{n_A}\right)^{1-\gamma} + \frac{1-\alpha}{w_M}A_M\left(A_A\left(\frac{l}{n_A}\right)^{1-\gamma} - \underline{c}\right) + (1-\alpha)\underline{c}\right]^{\varphi_d} + h(1-n_A) \tag{E3}$$

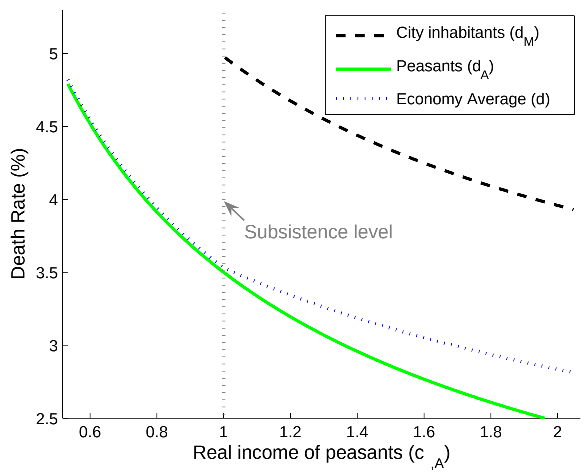
$$d_M = d_0\left[\alpha w_M + \frac{1-\alpha}{w_M}A_M\left(w_M - \underline{c}\right) + (1-\alpha)\underline{c}\right]^{\varphi_d} + \triangle d_M + h(1-n_A) \tag{E4}$$

The last term in (E3) and (E4) represents the Horsemen effect as a function of the urbanization rate $n_M = 1 - n_A$. For a given population size N we now have a system of 4 equations [(E1)-(E4)] and 4 unknowns (n_A, w_M, d_A, and d_M) that we solve numerically. Given these variables, it is straightforward to calculate the urbanization rate n_M, rural wages w_A from (2.4), and workplace-specific real income (or consumption) $c_{\bullet,i}$ from (2.9). Finally, workplace-specific birth rates are given by (2.10).

All calculations up to now have been for a given N. For small initial population, births outweigh deaths and N grows until diminishing returns bring down p.c. income enough for $b = d$ to hold. The opposite is true for large initial N. To find the long-run equilibrium with constant population, we derive b and d from (2.13) and (2.14). We then iterate the above system of equations, deriving N_t in each period t from (2.15), until the birth and death schedules intersect (point B in Figure 3.3). The long-run equilibrium level of population depends on the productivity parameters A_A and A_M, and on the available arable surface, L.

Under unconstrained migration, expected utility in each period is identical for peasants and manufacturing workers. Rural and urban population is given by $N_A = n_A N$ and $N_M = n_M N$, respectively. But is there migration in the long-run equilibrium? To answer this question, Figure 2.4 shows the workplace-specific death rates as a function of real income.[30] We calibrate city excess mortality including the effects of war and trade as $\triangle d_M = 1.5\%$. Equilibrium death rates in cities are higher than in the countryside.[31] Birth rates, on the other hand, are similar in both workplaces.[32] With stagnant total population and no migration, N_M would therefore decline continuously. This implication of our model is in line with the finding in the historical overview section that early modern European cities would have disappeared without a constant inflow of population.

Figure 2.4: Death rates by workplace and average death rate



Congestion and Constrained Migration

Major changes in the urban-rural income differential provide substantial incentives for migration. However, the short-term capacity of cities to absorb migrants is limited because new dwellings and infrastructure must be provided. Building new dwellings and urban infrastructure was one of the costliest undertakings in the early modern economy. Too many migrants caused congestion, making further movement to cities unattractive. In the interest of simplicity, we capture congestions effects with an upper limit

[30] As in Figure 3.2, we use peasants' consumption to represent real income. Urban income is a multiple of rural income, as implied by (2.8). Each point on the horizontal axis therefore corresponds to an urban real income level $c_{\bullet,M} > c_{\bullet,A}$. Note that for $c_{\bullet,A} < \underline{c}$, urban death rates are not defined since all individuals work in the countryside.

[31] The higher real income of manufacturing workers drives down d_M according to (2.11). However, this income effect is overcompensated by the higher background mortality in cities $\triangle d_M$.

[32] With an urban wage premium (relative to subsistence) in the range of 30% and birth rate elasticity $\varphi_b = 1.41$, as in our baseline calibration, b_M and b_A deviate by less than 0.05%.

to the growth rate of cities, $\overline{\nu}$.[33] When shocks are large, implying large wage differentials, the migration constraint becomes binding. It then takes time until the population shares reach their long-run equilibrium levels n_A^{LR} and n_M^{LR}.

Let $N^\star_{A,t}$ and $N^\star_{M,t}$ be the number of individuals living in the countryside and cities, respectively, at the beginning of period t *before* migration occurs. This 'native' population is determined by workplace-specific fertility and mortality in the previous period:

$$N^\star_{i,t} = (1 + b_{i,t-1} - d_{i,t-1})\, N_{i,t-1} \tag{2.19}$$

where $N_{i,t-1}$ is the number of agents that live at workplace $i = \{A, M\}$ during period $t-1$, *after* migration has taken place. Let M_t^u be the level of migration necessary to (immediately) establish long-run population levels $N_i^{LR} = n_i^{LR} N$ in period t, i.e., the migration that would take place if it were unconstrained:

$$M_t^u = N^\star_{A,t} - N_A^{LR} = N_M^{LR} - N^\star_{M,t} \tag{2.20}$$

There are two ways to calculate M_t^u, since migration out of agriculture (first term in (2.20)) must equal migration into cities (second term). M_t^u is positive if migration goes from the countryside to cities, i.e., if the number of native peasants is larger than the optimal long-run rural population, and negative if migration takes the opposite direction. Next we derive the growth of city population that occurs when migration is unconstrained, reaching the long-run equilibrium instantaneously.

$$\nu_t \equiv \frac{M_t^u}{N^\star_{M,t}} = \frac{N_M^{LR} - N^\star_{M,t}}{N^\star_{M,t}} = \frac{n_M^{LR} - n^\star_{M,t}}{n^\star_{M,t}} \tag{2.21}$$

As this equation shows, the growth rate of city population is equal to the growth of the urbanization rate – a fact that we will use to calibrate $\overline{\nu}$. The magnitude of M_t^u, and thus the likelihood that congestion constrains migration, is the larger the more the long-run population distribution deviates from actual values. If ν_t exceeds the upper bound for into-city migration, the constraint $\overline{\nu}$ becomes binding. The number of migrants under this constraint is given by $M_t^c = \overline{\nu} N^\star_{M,t}$, that is, urban population grows at the rate $\overline{\nu}$. Together with (2.19), this gives the law of motion for workplace-specific population under constrained migration.

$$\begin{aligned} N_{A,t} &= N^\star_{A,t} - \overline{\nu} N^\star_{M,t} \\ N_{M,t} &= N^\star_{M,t} + \overline{\nu} N^\star_{M,t} \end{aligned} \tag{2.22}$$

The agricultural workforce in period t is thus composed of rural offspring and surviving peasants from $t-1$, less the ones migrating to cities (until congestion makes these places unattractive). The city population consists of surviving manufacturing workers and urban offsprings, augmented by the migrants from the countryside.

The equilibrium values of wages, prices, and income under constrained migration are derived from the location-specific workforce given by (2.22). Rural wages are obtained directly from (2.4), while (E1) can be re-arranged to recover urban wages:

$$w_M = \frac{1-\alpha}{\alpha(1-n_A)} \left[A_A n_A^\gamma l^{1-\gamma} - \underline{c} \right] \tag{2.23}$$

[33] Migration in the opposite direction plays no role in our model because any income increase (following the Plague or technological progress) makes cities more attractive than the countryside via the manufacturing demand channel.

Manufacturing products are sold at $p_M = w_M/A_M$. Workplace specific real income, fertility, and mortality are then calculated from (2.9), (2.10), and (2.11), respectively.

2.4 Calibration and Simulation Results

In this section we explain the calibration of our model and simulate it with and without the Horsemen effect. We choose parameters in order to match historically observed fertility, mortality, and urbanization rates in early modern Europe. We then simulate the impact of the plague and derive the long-run levels of p.c. income and urbanization in the centuries following the Black Death. Finally, we add to our model the alleviating effect on birth rates that the European Marriage Pattern provided.

2.4.1 Calibration

In order to calibrate our model, we follow the procedure outlined in section 2.3.5: The intersection of birth and death schedule determines per-capita income and equilibrium population size. Urbanization rates in Europe before the Black Death were about 2.5%.[34] For cities to exist in our model, we need above-subsistence real income (and consumption) $c_{\bullet,i} > \underline{c}$ in the long-run pre-plague equilibrium. For the intersection of b and d to lie to the right of $\underline{c}$, we must have death rates higher than birth rates at the subsistence level, i.e., $d_0 > b_0$ in equations (2.10) and (2.11). Kelly (2005) estimates the elasticity of death rates with respect to income, using weather shocks as exogenous variation. We use his estimate for England over the period 1541-1700, $\varphi_b = -0.55$, as a best-guess for Europe. Regarding the elasticity of birth rates with respect to real income, we use his estimate of $\varphi_b = 1.41$ for Europe.[35] Regarding the level of birth and death rates, we use 3.5% in the pre-plague equilibrium, corresponding to the cumulative birth rates reported by Anderson and Lee (2002). This, together with the elasticities and the equilibrium urbanization rate of 2.5%, implies $b_0 = 3.2\%$ and $d_0 = 3.5\%$. As discussed in the historical overview section, we estimate that death rates in European cities were approximately 50% higher than in the countryside. This implies a (conservative) value of $\triangle d_M = 1.5\%$.

Scale does not matter in our model. Solely the productivity parameters A_A and A_M, together with the land-labor ratio l determine individual income. Thus, for any equilibrium p.c. income derived from the intersection of b and d, we can recover the corresponding population N.[36] We choose parameters such that initial population is unity ($N_0 = 1$). This involves $A_A = 0.460$, $A_M = 0.535$, and $L = 8$, where land is fixed such that its hypothetical rental rate is 5%.[37] Our calibration also implies the desired urbanization rate $n_{M,0} = 2.5\%$ and a price of manufacturing goods that is double the price of agriculture

[34] Maddison (2003) reports 0% in 1000 and 6.1% in 1500; DeVries (1984) documents 5.6% in 1500. Our 2.5% for the 14th century is at the upper end of what we expect, given that wages stagnated throughout the millenium before the plague. We deliberately make this conservative choice, leaving less urbanization to be explained by our story.

[35] This number is bigger than the estimates in, say, Crafts and Mills 2007, or in Lee and Anderson 2002. Because of the important endogeneity issues in deriving any slope coefficient, the IV-approach by Kelly is more likely to pin down approximate magnitudes than identification through VARs or through Kalman filtering techniques.

[36] For example, rural population is implicitly given by (2.4), and is the larger (for a given wage w_A) the more land is available. We calculate the long-run equilibrium by solving the system (E1)-(E4) and iterating over population until $b = d$. This procedure gives the long-run stable population as a function of fertility and mortality parameters, productivity, and land area.

[37] Recall that we assume no property rights to land. The size of L is therefore not important for our results – it could also be normalized to one and included in A_A. We leave L in the equations for the sake of arguments involving the land-labor ratio.

products, i.e., $p_M = 2$.[38]

For the baseline model, we calibrate the parameters γ to fit the average historical labor share in agriculture, using data for England over the period 1700-1850, which implies $\gamma = 0.6$. This corresponds to the land income share of 40% suggested by Crafts (1985), and is almost identical with the average in Stokey's (2001) two calibrations. We normalize the minimum food consumption $\underline{c}$ to unity. For low income levels, all expenditure goes to agriculture. With higher productivity, manufacturing expenditure share and urbanization grow in tandem. To derive this relationship, we pair income data from Maddison (2007) with urbanization rates from DeVries (1984). In the model, the responsiveness of urbanization to income is governed by the parameter α. The data for Europe show that the urbanization rate rose from 5 to 10 percent between 1500 and 1800, while p.c. income grew by 50%. The corresponding model parameter that approximately reflects this relationship is $\alpha = 0.6$. Figure 3.2 is derived using this value.

In the centuries before 1700, labor productivity grew at an average rate of roughly 0.05-0.15% per year (Galor 2005). We use an exogenous growth rate of agricultural and manufacturing TFP, A_A and A_M, of $\tau = 0.1\%$ in our simulations with technological progress. In order to quantify the upper bound for city growth, reflecting congestion in our model, we use DeVries' (1984) urbanization data for 1500-1800. The largest observed growth rate of urbanization in Europe over this period is $\overline{\nu} = 0.38\%$ between 1550 and 1600.

After the Black Death, the Horsemen effect comes into play. Means for warfare and trade grow with p.c. income, and the increased mobility leads to an ongoing dispersion of germs. According to equation (2.12), the Horsemen are at work when the urbanization rate n_M is larger than the threshold level n_M^h. We choose $n_M^h = 2.5\%$, corresponding to the pre-plague urbanization rate. Therefore, the Horsemen effect begins to work its wonders immediately after the Black Death, though not yet with full force. The effect is linearly increasing in the urbanization rate until reaching its maximum. In order to calibrate the maximum impact of the Horsemen channel on mortality, we use data on war-related deaths and epidemics from Levy (1983). His data show that, in a typical year, more than one European war was in progress – there were 443 war years during the period 1500-1800, normally involving three or more powers. Since it is the movement of armies, and not just military engagements that caused death, we count the territories of combatant nations as affected if they were the locus of troop movements. The weighted average produces a war-related effect of an additional 0.75% deaths per annum. To this we add a guestimate of 0.25%. This is motivated by the spread of disease through additional trade, also facilitated and encouraged by the wealth of cities – few of the goods on the plague ship in Marseille harbor in 1720 would have carried goods for the consumption of peasants. Overall, our best guess for the maximum size of the Horsemen effect is thus $h_{\max} = 1\%$. This value is reached in the first half of the 17th century. War frequency was almost double what it had been a century before, and the devastation wrought by the Thirties Years War was the most severe in any armed conflict until the 20th century (Levy 1982). Urbanization rates reached 8% at this time (De Vries, 1984). The implied slope parameter of the Horsemen function is therefore $\delta = h_{\max}/(0.08 - n_M^h) = 1.82$. Table 2.1 summarizes the calibration parameters.

[38] Other values of this parameter, resulting from different A_M relative to A_A, do not change our results.

Table 2.1: Baseline Calibration

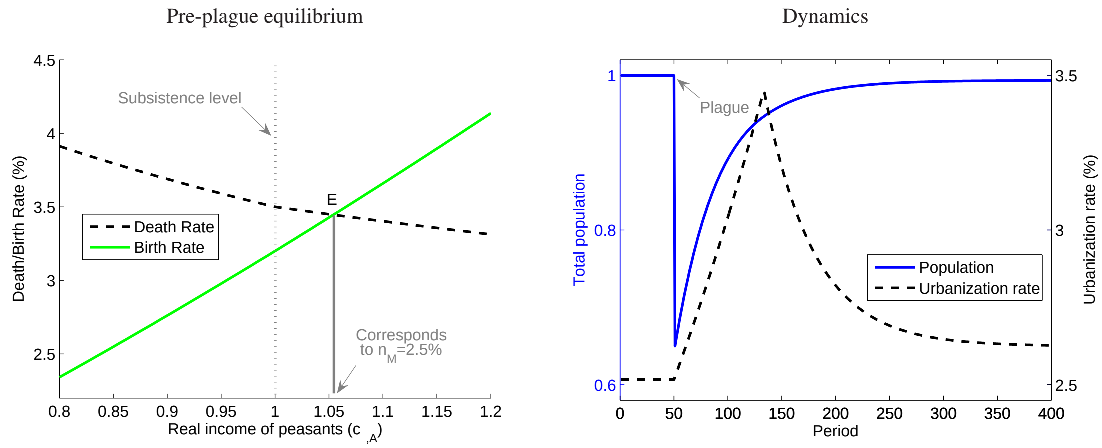
Symbol	Interpretation	Value
Parameters		
α	Food expenditure share (as income $\rightarrow \infty$)	0.6
γ	Labor share in agriculture	0.6
$\underline{c}$	Subsistence food consumption	1
L	Land	8
A_A	Agriculture technology parameter	0.460
A_M	Manufacturing technology parameter	0.535
τ	Rate of technological progress	0.1%
b_0	Birth rate at $c = \underline{c}$	0.032
d_0	Death rate at $c = \underline{c}$	0.035
φ_b	Elasticity of birth rates wrt. income	1.41
φ_d	Elasticity of death rates wrt. income	-0.55
$\triangle d_M$	City excess mortality	0.015
$h_{\max}$	Maximum Horsemen effect	0.01
n_M^h	Threshold for Horsemen effect	0.025
δ	Slope parameter for Horsemen effect	1.82
$\overline{\nu}$	Upper bound on city growth due to congestion	0.0038
Resulting values in long-run equilibrium before Black Death		
N_0	Population	1.00
$n_{A,0}$	Urbanization rate	2.5%
$b_0 = d_0$	Economy-average birth and death rate	3.5%
$p_{M,0}$	Relative price of manufacturing goods	2.00

2.4.2 Plague and Equilibrium without Horsemen Effect

The left panel of figure 2.5 shows the pre-plague long-run equilibrium corresponding to our baseline calibration. The fertility and mortality schedule intersect at 3.5%, while 2.5% of the population live in cities. The economy is trapped in Malthusian stagnation in point E. One-time increases in productivity lead to higher income and therefore population growth. As a consequence, the land-labor ratio falls and drives per-capita income back to its long-run equilibrium value.

The effect of a one-time technological improvement on p.c. income is very similar to the impact of the plague in our model: While the former raises TFP, the latter increases the land-labor ratio; both result in higher wages, according to (2.4). The right panel of figure 4 shows the effect of the Black Death when all model parameters are unchanged. Before the plague, population and urbanization rate stagnate in the absence of technological progress. The Black Death kills one third of the population, similar to the devastation documented in 14$^{\text{th}}$ century Europe. As an immediate consequence, wages, p.c. consumption, and urbanization rates rise. In the aftermath of the plague, population grows because the economy is now situated to the right of the long-run equilibrium in point E, such that fertility outweighs mortality. The falling land-labor ratio then drives the economy back to E, where all variables are back at their pre-plague

Figure 2.5: Long-run impact of the plague, ceteris paribus

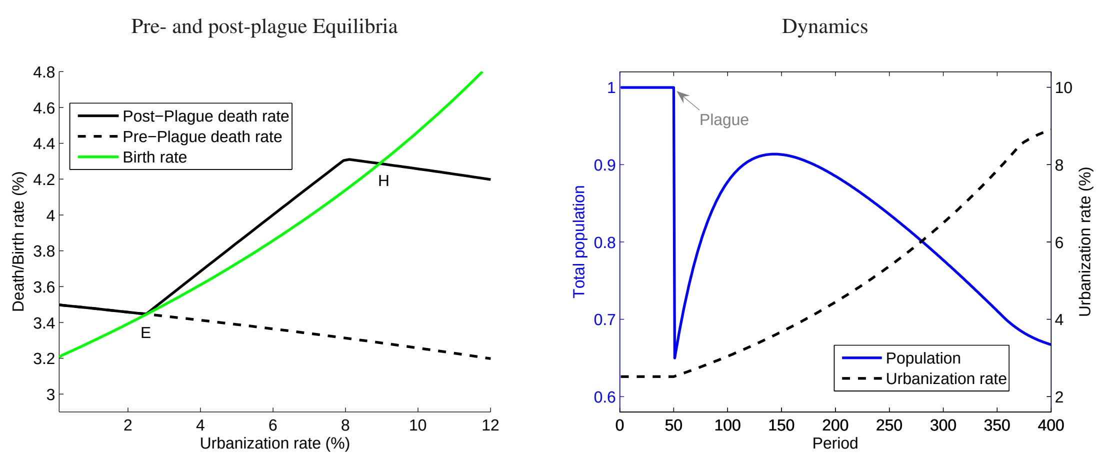


values. Things look different in the presence of the Horsemen effect.

2.4.3 Long-run Equilibrium with Horsemen Effect

Following the plague in Europe, higher p.c. income increases trade, but also means for warfare. The enhanced mobility constantly spreads epidemics and therefore raises country-wide mortality. The size of this Horsemen effect grows with urbanization, as given in (2.12). The left panel of figure 5 shows that the equilibrium with the Horsemen effect (point H) involves higher birth and death rates (about 4%), more p.c. consumption, and higher urbanization. Point H is a unique and stable equilibrium where all variables are in a stalemate in the absence of technological progress. The economy converges to this equilibrium in the aftermath of the Black Death (right panel of figure 5). Surviving individuals and their descendants are therefore better off then their ancestors before the plague. Famously, it took until the 19th century for wages to recover the level last seen at the post-plague peak (Clark 2005).

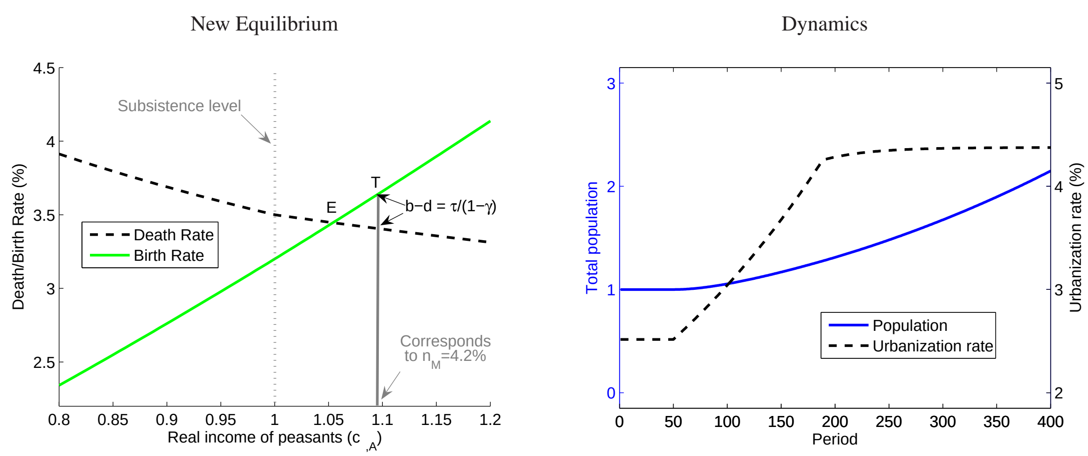
Figure 2.6: Impact of the plague with Horsemen effect



2.4.4 Technological Progress and Model Fit

Technological progress in premodern times alone is not enough to escape from the Malthusian trap. While the growing population completely eats up the fruits of one-time inventions, *ongoing* progress implies higher, but still stagnating, long-run p.c. income. Technology constantly improves p.c. income in this case, and population growth responds, offsetting any gains. This corresponds to a long-run equilibrium in point T in figure 7, where the birth rate exceeds the death rate and technological progress is exactly offset by the falling land-labor ratio.[39] The right panel of figure 7 illustrates the orders of magnitude involved. The rate of technological change before the Industrial Revolution was low, approximately 0.1% (Galor 2005). For purposes of illustration, progress is assumed to set in after 50 periods of stagnating technology. As the figure shows, this raises the urbanization rate by less than 2%. Note that this is an extreme scenario where the economy jumps from complete stagnation to continuous inventions. The corresponding increase of urbanization is thus an upper bound for the impact of technology on individual income. Therefore, technological progress cannot be a candidate to explain the rise of Europe in the early modern period.

Figure 2.7: Effect of ongoing technological progress



Next, we investigate the fit of our model, including the Horsemen effect and the observed rate of technological progress. While the former alone can account for almost all the observed increase in European urbanization (see figure 2.4.3), the latter helps to explain the growing population. Figure 2.8 shows the corresponding simulation results together with the data. Our model performs well in reproducing both population growth and urbanization.

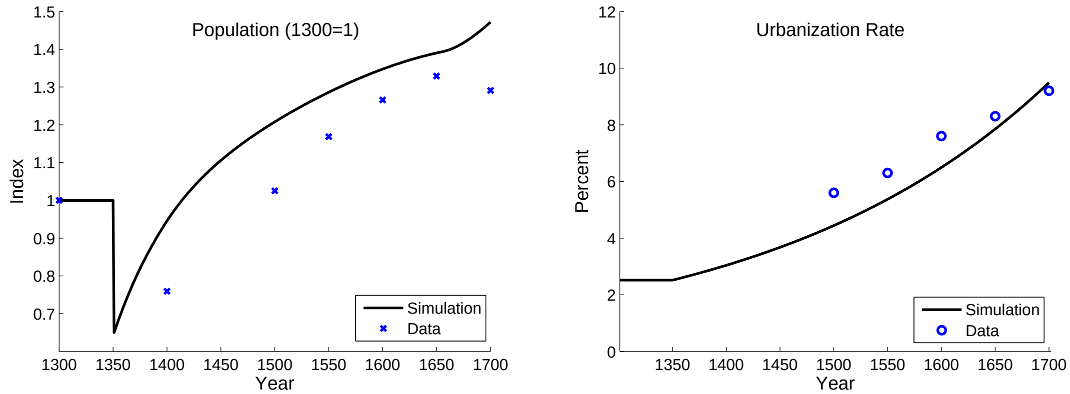
2.5 Extensions

2.5.1 The European Marriage Pattern

Europe curtailed fertility in an important way. In a normal Malthusian setting, fertility should have eroded all gains in living standards quickly. Lower birth rates for a given income have a similar effect as higher

[39] Equation (2.3) with constant p.c. income (and thus constant agricultural labor share) implies that the population grows at the rate $\tau/(1-\gamma)$ in the long-run equilibrium.

Figure 2.8: Europe: Simulation Results vs. Data



Data sources: Population from Kremer (1993). Urbanization rates from DeVries (1984).

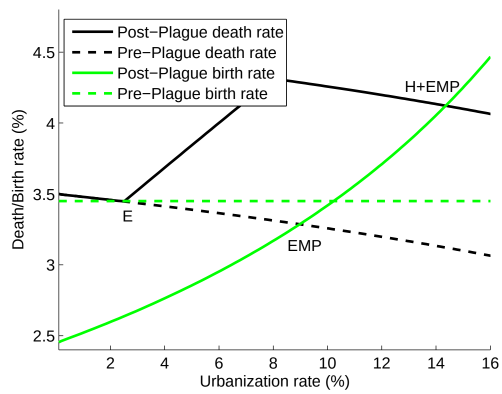
background mortality: both alleviate population pressure on the land-labor ratio. Average realized fertility rates in early modern Europe were approximately equal to those in China, despite markedly higher living standards (Clark 2007). At Chinese levels of per capita income, European fertility would be much lower because of fertility restriction. In figure 2.9, 17$^{\text{th}}$ century China would be close to point E, implying that English fertility with the EMP would have been 0.75% below the corresponding value for China. We do not know when the European marriage pattern emerged. Some authors have argued that the plague was critical (Van Zanden and deMoor 2007). If so, then some of the increase in European incomes after 1350 has to be attributed to the plague's impact on fertility.

Births in England were probably unusually responsive to economic conditions (Lee 1981, Wrigley and Schofield 1981). England was also ahead of the European average in terms of its income and city growth: p.c. income grew by 75% between 1500 and 1700 (50% in Europe), and urbanization rates more than quadrupled from 3% to over 13% (and 20% in 1800) (Maddison 2007, DeVries 1984). We now turn to investigating the contribution of the EMP to income and city growth in our model. We suppose that birth rates are not responding to income before the Black Death, so that $\varphi_b^{\text{before}} = 0$. After the plague the EMP emerges, shifting the birth schedule downwards by 1% (corresponding to $0.3 \times 3.5\%$) and making it responsive to income such that $\varphi_b^{\text{after}} = 1.4$, as in the baseline calibration.[40]

Figure 2.9 shows the EMP simulation results. In the absence of the Horsemen effect, the shift and turn of the birth schedule move the economy from the pre-plague equilibrium E to the EMP equilibrium. Part of the downward-shift is compensated by the positive response of birth rates to growing income after the plague. The Horsemen effect creates an additional rise in urbanization (equilibrium H+EMP). Instead of an urbanization rate of 2.5%, we predict a rate of 9% based on the rotated fertility schedule. The rise from 9 to 15% is due to the Horsemen effect. Both effects appear to be equally important, and together they match the observed increase in England's urbanization rate. This underlines the importance of the Horsemen effect for increasing living standards in early modern Europe – in continental Europe, where the EMP was weaker, the Horsemen contribution was likely even more significant.

[40]The corresponding parameter values are $b_0^{\text{before}} = 0.0345$ and $b_0^{\text{after}} = 0.0245$. With all other parameters unchanged, this implies an equilibrium urbanization rate of 2.5% before the plague.

Figure 2.9: England: Pre-and post-plague equilibria with EMP and Horsemen



The fact that the outward shift of the death schedule did not lead to a much sharper decline in population is also a result of the European Marriage Pattern. Had it not been for the sharp response of fertility due to higher living standards, the reduction in population pressure after 1350 would have been even stronger.

2.6 Concluding Remarks

Standard accounts of the transition from "Malthus to Solow" emphasize the near-stability of incomes before 1800 (Hansen and Prescott 2002). While sensible compared to modern rates of growth, it is incorrect by historical standards. The Dutch Republic and England in 1700 had per capita incomes that were extraordinary compared to all ages that came before, and contemporaries saw them as such (DeVries 1976). This precocious rise in incomes long before industrialization may have been an important factor contributing to the ultimate economic, military and political ascendancy of Western Europe from the 19th century onwards.

In this paper, we argue that a simple two-sector extension of the standard Malthusian model can shed new light on the puzzling rise of European per capita incomes. Many interpretations of the "rise of Europe" emphasize high rates of innovation, compared to Asia (Mokyr 1990), or fertility restriction (Wrigley 1988). We argue that, in a Malthusian setting, better technology cannot explain the "Great Divergence", and we also show that fertility restriction alone is insufficient. Instead, we build a model in which per capita living standards can rise markedly without technological change or fertility decline. Many unified growth models generate the early transition from stagnation to sustained growth by means of a delayed response of fertility to wages. This allows per capita incomes to rise slowly but steadily in tandem with population. We argue that this cannot be realistic in most settings, because fertility responds "too rapidly" to permit anything other than a short-lived increase in living standards. In a micro-founded model, we show that only very large, negative shocks can be followed by a marked delay between rising incomes and return to earlier population levels. We argue that the Black Death hitting Europe in the 14th century was precisely such a shock, lifting wages and per capita incomes for several generations. Richer individuals began to demand more urban goods, and because early modern European cities were "graveyards" (Bairoch 1991), incomes could permanently exceed conventionally-measured subsistence levels. This is particularly true because city growth acted as a catalyst for European belligerence and the spread of disease through trade – a link we call the "Horsemen of Growth".

One implication of our results is that urbanization is not simply an indicator of higher levels of development, as assumed in some recent work (Acemoglu, Johnson and Robinson 2005). City growth also provided a mechanism that made higher per capita incomes sustainable in a Malthusian setting. Our paper has emphasized the contrast between early modern Europe and the rest of the world. Future research should examine if the model developed in this paper can also explain the growing differences between Northwestern and Southern Europe. Did differences in political structure allow the self-sustaining rises in incomes to persist for longer in the North? Did greater preferences for urban goods, or differences in sanitary practices drive up mortality rates differently? While many histories of the "rise of the West" have been written from a technological perspective, we argue that differences in mortality (and also, fertility) were far more potent determinants of pre-industrial living standards.

References

[1] Acemoglu, Daron, Simon Johnson and James A. Robinson, 2005, 'Rise of Europe: Atlantic Trade, Institutional Change and Economic Growth', American Economic Review 95: 546-579.

[2] Anderson, Michael; and Ronald Lee 2002, 'Malthus in state space: Macro economic-demographic relations in English history, 1540 to 1870', Journal of Population Economics 15: 195-220.

[3] Allen, Robert C., 2001, 'The Great Divergence in European Wages and Prices from the Middle Ages to the First World War', Explorations in Economic History 38: 411-447.

[4] Allen, Robert C., 2004, 'Agricultural Productivity and Rural Incomes in England and the Yangtze Delta, c. 1620-c.1820', Nuffield working paper.

[5] Allen, Robert C., Tommy Bengtsson, and Martin Dribe, eds., 2005, Living Standards in the Past: New Perspectives on Well-Being in Asia and Europe, Oxford, OUP.

[6] Bairoch, Paul, 1991, Cities and Economic Development from the Dawn of History to the Present, Chicago: University of Chicago Press.

[7] Bosworth, Barry and Susan Collins, 2003, 'The Empirics of Growth: An Update', Brookings Papers on Economic Activity.

[8] Brainerd, Elizabeth, and Mark Siegler, 2003, 'The Economic Effects of the 1918 Influenza Epidemic', Williams College typescript.

[9] Brewer, John, 1991, The Sinews of Power, Cambridge, MA: Harvard University Press.

[10] Broadberry, Stephen and Bisnupriya Gupta, 2006, 'The Early Modern Great Divergence: Wages, Prices and Economic Development in Europe and Asia, 1500-1800', Economic History Review 59: 2-31.

[11] Clark, Gregory, 2005, 'The Condition of the Working Class in England, 1209-2003', Journal of Political Economy 113: 1307-1340.

[12] Clark, Gregory, 2007, Farewell to Alms, Princeton: PUP.

[13] Crafts, NFR, and Terence Mills, 2007 (forthcoming), 'From Malthus to Solow: How the Malthusian Economy Really Evolved', Journal of Macroeconomics.

[14] Coleman, D.C., 1983, 'Proto-Industrialization: A Concept Too Many', The Economic History Review 36 (3): 435-448.

[15] DeVries, Jan, 1976, "The Economy of Europe in an Age of Crisis, 1600-1750," Cambridge: CUP.

[16] DeVries, Jan, 1984, "European Urbanization 1500-1800". Methuen and Co Ltd. London.

[17] Diamond, Jared, 1997, Guns, Germs, and Steel: The Fates of Human Society. Norton: New York.

[18] Galley, Chris, 1998, The Demography of Early Modern Towns: York in the Sixteenth and Seventeenth Centuries, Liverpool: Liverpool University Press.

[19] Galor, Oded and David Weil, 2000, ’Population, Technology and Growth: From the Malthusian Regime to the Demographic Transition and Beyond,’ American Economic Review 90: 806-828.

[20] Galor, Oded, 2005, ’From Stagnation to Growth: Unified Growth Theory’, in Philip Aghion & Steven Durlauf, eds., Handbook of Economic Growth, Vol. 1A. Amsterdam: North-Holland.

[21] Glick, Reuven and Alan M. Taylor, 2006, ’Collateral Damage: Trade Disruption and the Economic Impact of War’, UC Davis working paper.

[22] Hanley, Susan B., 1997, Everyday Things in Premodern Japan: The Hidden Legacy of Material Culture, Berkeley: University of California Press.

[23] Hajnal, John, 1965, ’European Marriage Pattern in Historical Perspective’, in D.V. Glass and D.E.C. Eversley, eds., Population in History, London: Arnold.

[24] Hansen, Gary and Edward Prescott, 2002, ’Malthus to Solow’, American Economic Review 92 (4): 1205-1217.

[25] Hayami, Akira, 2001, The Historical Demography of Pre-Modem Japan, Tokyo: University of Tokyo Press.

[26] Hess, Gregory, 2003, ’The Economic Welfare Cost of Conflict: An Empirical Assessment’, CESifo Working Paper no. 852.

[27] Hoeffler, Anke, and Marta Reynal-Querol, 2003, ’Measuring the Costs of Conflict’, Oxford University working paper.

[28] Hughes, Robert, 1992, Barcelona, Vintage: New York.

[29] Kelly, Morgan, 2005, ’Living Standards and Population Growth: Malthus was Right.’, UC Dublin working paper.

[30] Lagerlöf, Nils-Petter, 2003, ’From Malthus to Modern Growth: Can Epidemics Explain the Three Regimes?’, International Economic Review 44 (2): 755-777.

[31] Landers, John, 2003, The Field and the Forge, Oxford: OUP.

[32] Lee, Ronald, and Michael Anderson (2002) “Malthus in State Space: Macro Economic-Demographic Relations in English History,” Journal of Population Economics, 15 (2): 195-220.

[33] Lee, Ronald, 1981, "Short-term Variation: Vital Rates, Prices, and Weather," in E. Anthony Wrigley and Roger S. Schofield, The Population History of England 1541—1871, Cambridge, CUP: 356—401.

[34] Levy, J., 1983, War in the Modern Great Power System, 1495-1975, The University Press of Kentucky, Lexington, Kentucky.

[35] Lindegren, Jan, 2000, ’Men, Money and Means’, in P. Contamine, ed., The Origins of the Modenr State in Europe, Oxford: OUP.

[36] Liu, Ts'ui-Jung, 1990, 'Demographic Aspects of Urbanization in the Lower Yangzi Region of China, c. 1500-1900,' in Ad van der Woude, Jan de Vries, and Akira Hayami, eds., Urbanization in History, Oxford: Clarendon Press: 328-351.

[37] Macfarlane, Alan, 1997, The Savage Wars of Peace: England, Japan and the Malthusian Trap. Oxford: Basil Blackwell.

[38] McNeill, William H., 1977, Plagues and People, Anchor: New York.

[39] Maddison, Angus, 2003, The World Economy. A Millennial Perspective, Paris: OECD.

[40] Maddison, Angus, 2007, Historical Statistics. University of Groningen webpage: http://www.ggdc.net/maddison/.

[41] Mokyr, Joel, 1990, The Lever of Riches, Oxford: OUP.

[42] Mullett, Charles, 1936, 'The English Plague Scare of 1720-23', Osiris 2: 484-516.

[43] Murdoch, James, and Todd Sandler, 2002, 'Economic Growth, Civil Wars, and Spatial Spillovers', Journal of Conflict Resolution 46 (1): 91-110.

[44] Ogilvie, Sheilagh, and Markus Cerman, eds., 1996, European Proto-industrialization, Cambridge: CUP.

[45] Pomeranz, Kenneth, 2000. The Great Divergence: China, Europe, and the Making of the Modern World, Princeton: PUP.

[46] Rozman, Gilbert, 1973, Urban Networks in Ch'ing China and Tokugawa Japan, Princeton: PUP.

[47] Sala-i-Martin, Xavier, Gernot Doppelhofer, and R Miller, 'Determinants of Long-Term Growth: A Bayesian Averaging of Classical Estimates (BACE) Approach', American Economic Review 94 (4): 813-835.

[48] Slack, Paul, 1981, 'The Disappearance of the Plague: An Alternative View', The Economic History Review 34 (3): 469-76.

[49] Smith, Adam, 1776 (1976), Wealth of Nations.

[50] Tilly, Charles, 1990, Coercion, Capital, and European States, AD 990-1992, Oxford: Blackwells.

[51] Voigtländer, Nico and Hans-Joachim Voth, 2006, 'Why England? Demographic Factors, Structural Change and Physical Capital Accumulation during the Industrial Revolution', Journal of Economic Growth 11: 319-361.

[52] Weil, David, 2004, Economic Growth, New York: Addison-Wesley.

[53] Wells, Herbert G., 1905, A Modern Utopia.

[54] Woods, Robert, 2003, 'Urban-Rural Mortality Differentials: An Unresolved Debate', Population and Development Review 29 (1): 29-46.

[55] Young, Alwyn, 2005, 'The Gift of the Dying: The Tragedy of AIDS and the Welfare of Future African Generations,' Quarterly Journal of Economics 120: 243-266.

[56] Van Zanden, Jan Luiten and Tine de Moor, 2007, 'Girlpower. The European Marriage Pattern (EMP) and Labour Markets in the North Sea Region in the Late Medieval and Early Modern Period', unpublished ms.

[57] Wrigely, E.A., 1988, Continuity, Chance, and Change, Cambridge: CUP.

[58] Wrigley, E.A. and Roger S. Schofield, 1981, The Population History of England 1541—1871, Cambridge, CUP.

Chapter 3

Many Sectors Meet More Skills: Intersectoral Linkages and the Skill Bias of Technology

3.1 Introduction

This paper shows that skill upgrading in one sector increases skill demand in many other sectors, because of linkages operating through the use of intermediate products. This channel has been ignored by the literature so far, despite the fact that more than half of a final product's value is embedded in intermediates. I construct a measure of input-embedded skills, matching input-output tables with workforce data for detailed U.S. manufacturing sectors. Input skill intensity is defined as the weighted average share of white-collar workers employed in the production of a sector's intermediate inputs.[1] Figure 3.1 presents a novel stylized fact: A strong positive correlation between input skill intensity and skills employed in final production in U.S. manufacturing.[2] I argue that this finding implies an intersectoral technology-skill complementarity (ITSC): Skills used in intermediate production are complementary to skills required in the further processing of intermediates or their integration into redesigned final products. Using an estimation strategy derived from a labor-demand framework, I show that the ITSC is quantitatively important, explaining more than one third of the increase in white-collar labor demand in U.S. manufacturing.

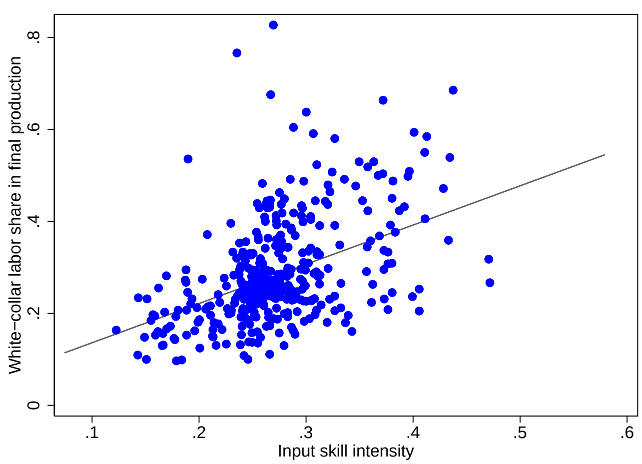


Figure 3.1: Skilled labor share in final production vs. input skill intensity

Notes: Data are for 358 U.S. manufacturing sectors in 1992. Input skill intensity is calculated as the weighted average share of white-collar workers employed in the production of a sector's inputs. Only inputs purchased outside a sector are taken into account. See Section 3.3 for a formal description and data sources, and Section 3.4 for regression results and robustness.

[1] White-collar workers – including personnel engaged in supervision, installation and servicing, professional, technological, and administrative – have been widely used to proxy for skilled labor. See in particular Berman, Bound, and Griliches (1994).

[2] The figure presents cross-sectional observations in 1992. The correlation is very similar for any other benchmark year (5-year intervals) between 1967 and 1992.

Empirical evidence suggests that this complementarity works through product innovation. Building on the seminal work by Nelson and Phelps (1966), much econometric and case-study evidence indicates that highly skilled workers are not merely more productive, but are also good innovators, adapt better to technological change, and speed the process of technological diffusion [Bartel and Lichtenberg 1987; Goldin and Katz 1998; Doms, Dunne and Troske 1997]. Because of this innovation-skill complementarity, an upstream supplier employing highly educated workers turns out innovative intermediates. Upstream product improvement induces innovation at the downstream level, which in turn increases skill demand.[3] The argument does not depend on the direction of causality; the complementarity also works from producers to suppliers. An example is a cutting-edge downstream firm demanding innovative intermediate inputs, so that its upstream supplier needs highly skilled workers.

Because of the ITSC, inputs used in the production process are not only 'intermediate' in the standard semi-manufactured sense, but also 'intermediaries' that transmit skill requirements across industries. Therefore, product innovation affects labor demand not only in the corresponding firm or industry, as previous studies have argued, but also in other firms or sectors, via input-output linkages. These linkages deliver a multiplier that reinforces skill demand across firm and sector boundaries. For example, the invention and improvement of the transistor affected skill demand within and outside its sector of origin, the electronic components industry. Within this industry, the transistor enabled the production of more refined electronic parts, engineered by highly skilled workers. These innovative electronic components eventually became fundamental intermediate inputs for a large variety of other sectors, including computers, communication equipment, and controlling devices, where their integration went hand in hand with product innovation and skill upgrading. The improvement of these devices, in turn, enabled further innovation of electronic components. Innovation in the electronic components industry therefore drives skill demand in many other sectors. Eventually, it feeds back into the originating sector itself, creating a virtuous circle, or in effect a multiplier of skill upgrading.

This amplification mechanism closes an important gap in the empirical wage-inequality literature. While many variables have been shown to contribute to rising skill demand in a statistically significant way, accounting for the full scope has proved difficult. By adding the intermediate dimension, this paper shows that skill upgrading in one sector leads to rising skill demand in linked sectors, amplifying initial increases in skill demand along the production chain. It also suggests that a positive supply shock for skilled labor could lead to a rise in skill demand in the economy at large.

The empirical analysis in this paper is based on U.S. input-output data, paired with workforce characteristics from the NBER Manufacturing Industries Database at the detailed 4-digit level over the period 1967 to 1992. To quantify input skill intensity, I construct a variable that measures, for each sector, the proportion of white collar workers involved in the production of its intermediate inputs. The correlation between input skill intensity and skill share in final production is stable over time and robust to the inclusion of a large number of additional controls.[4] These results are confirmed when going beyond a mere

[3] As Scherer (1982) for the United States and Pavitt (1984) for Great Britain show, product innovation in upstream sectors serves to improve productivity and quality of output in the buying industries.

[4] These controls include sectoral fixed effects, time dummies, as well as various measures proposed in the wage inequality literature: capital intensity, shares of computer and high-tech capital, R&D intensity, and outsourcing. I also exclude inputs from similar industries in the calculation of input skill intensity in order to address the concern that common trends drive the observed correlation.

correlation analysis and using instruments to account for the endogeneity of input skill intensity and control variables. The estimated ITSC implies a multiplier of approximately 1.5 - 2. Consequently, an initial innovation (or shock) that causes immediate average skill upgrading of one percent translates into a total skill demand increase of up to two percent (for given relative wages), as innovations and skill demand spread across sectors, reinforcing each other.

Differentiated goods can be refined more readily than homogenous ones. Crude petroleum does not change, whether it is pumped out of the ground by laborers or university graduates, while the presence of engineers contributes to continuous improvement of electronic components. Therefore, differentiated inputs are more susceptive to 'skill embedding.' Innovations in the production of homogenous inputs improve processes rather than products, and thus have little effect on skill demand at the subsequent downstream level. I provide evidence for this assertion, combining data on sectoral product and process innovation from Scherer (1982) with Rauch's (1999) classification of product differentiation. The constructed cross-section shows that product innovation is more pronounced in sectors that produce differentiated goods. Thus, downstream users of differentiated intermediates purchase relatively more embedded product innovation. Next, I use Rauch's (1999) classification to construct a measure of input differentiation. I demonstrate that the ITSC is increasing in the degree of input differentiation, and is insignificant for sectors that use mainly homogenous inputs. In an additional analysis, I show that skill-intensive intermediates coincide with higher productivity only in those sectors that employ skilled workers able to handle them. In the absence of final production skills, skill-intensive intermediates could even harm output per worker. Thus, input-embedded and final production skills complement each other in raising productivity.

The rest of the paper is organized as follows. Section 3.2 reviews the related literature and presents a framework that incorporates complementarities between product innovation, skills embedded in intermediate inputs, and skilled labor in final production. Section 3.3 describes the data, explaining in detail the construction of my input skill intensity measure. Section 3.4 reports empirical results, documenting the intersectoral technology-skill complementarity, and confirms its robustness. In addition, I derive a regression from a labor-demand framework to estimate the ITSC's contribution to skill upgrading. I addresses endogeneity issues by using a set of IV regressions and check instrument validity, applying weak instrument tests and overidentifying restrictions. Section 3.5 integrates the novel stylized fact into the theoretical skill-biased technical change (SBTC) framework. I build a simple model that adds intermediate inputs to the standard SBTC setup. Therein, the relative productivity of skilled workers depends on the skills embedded in intermediates. A calibration exercise implies that the correlation of skill requirements along the production chain accounts for up to one half of SBTC in U.S. manufacturing over the sample period, underscoring the potential of my framework to reconcile key facts. Section 3.6 concludes.

3.2 Motivation and Framework

As the supply of skilled workers has risen, so has the skill premium. A large body of studies following Katz and Murphy (1992) documents substantial increases in wage inequality in the United States. Skill upgrading, i.e., a rise in skilled labor's share in employment and payroll, is also observed in other OECD countries as well as in developing countries.[5] Many explanations have been offered for the rising wage inequality, but two stand out: Trade liberalization and its effects on international patterns of specialization [e.g., Leamer 1996; Wood 1998; Feenstra and Hanson 1999], and skill-biased technical change – technological progress that shifts demand toward more highly skilled workers relative to the less skilled. Numerous studies quantify SBTC as a complementarity between capital (or technology) and skill, where computer-based information technologies (IT) play a central, although disputed role [DiNardo and Pischke 1997, Card and DiNardo 2002]. So far, the SBTC literature has treated technology-skill complementarities as a phenomenon within specific industries[6], within firms[7], and at the worker level[8], ignoring linkages across sectors. Some contributions add the role of complementarities among information technology, production organization, and product innovation [Milgrom and Roberts 1990] and link these to the observed increase in skill demand [Bresnahan et al. 2002].

Existing work can explain some of the rise in skill demand, but falls far short of accounting for all of it. The first prominent channel, international trade, has a between- and a within-industry component. The between-component, relocating production of low-skill-intensive industries to low-skill abundant countries, contributes little to the observed skill upgrading [Berman et al. 1994, Autor et al. 1998]. To explain the importance of observed within-industry skill upgrading, Feenstra and Hanson (1999) suggest outsourcing of low-skill intensive activities within firms or sectors. Their measure explains up to 15% of relative wage increases in U.S. manufacturing. The second group of explanations has used numerous variables to quantify the skill bias of technical change. Computers and other high-tech capital have been shown to contribute about 1/3 to the increase in white-collar labor demand in manufacturing [Feenstra and Hanson 1999, Autor et al. 1998].[9] The role of a broad capital-skill complementarity [Krusell, Ohanian, Ríus-Rull, and Violante 2000] has proved controversial. Finally, while studies document significantly positive coefficients on R&D intensity [Machin and van Reenen 1998, Autor et al. 1998], the variable itself changes relatively little over time. I show below that R&D intensity can account for about 5-10% of skill upgrading in manufacturing. All individual contributions together explain only about half of the overall magnitude.

Studies of SBTC have made the key (and limiting) assumption that complementarities are found at the individual worker, firm, or industry level. To this, I add complementarities across sectors, i.e., complementarities between input-embedded skills and skills employed in the subsequent processing of

[5] See Machin and van Reenen (1998), and Berman, Bound and Machin (1998) for evidence on the former; Pavcnik (2003), and Zhu (2005) on the latter group of countries.

[6] Berman et al. (1994) find that the rate of skill upgrading within U.S. manufacturing is strongly correlated with IT investment and R&D, and accounts for most of the demand shift towards skilled workers over the 1980s. This effect has been greater in more IT-intensive industries [Autor, Katz and Krueger 1998]. Autor, Levy and Murnane (2003) argue that computer capital substitutes for 'routine tasks' while it complements more complex 'nonroutine' tasks performed by skilled workers.

[7] Levy and Murnane (1996), Doms et al. (1997), and Bresnahan, Brynjolfsson and Hitt (2002) use broad measures of technological progress and provide evidence on firm and plant level skill-favoring demand shifts.

[8] Krueger (1993) and Autor et al. (1998) document a strong positive correlation between wages and computer use by workers. Epifani and Gancia (2006) point out scale increases as an additional channel for skill bias. See Bound and Johnson (1992) and Autor, Katz and Kearney (2008) for an assessment of alternative explanations of the observed relative wage changes. Katz and Autor (1999) and Sanders and Ter Weel (2000) summarize the literature at the three levels of aggregation.

[9] These estimates are to be interpreted with caution, as they take correlation coefficients as causal effects. Autor et al. (2003) investigate computer-induced task shifts in all sectors of the U.S. economy. Their approach can explain up to sixty percent of the relative demand shift favoring college labor, but half of this impact is due to task changes within nominally identical occupations. The remaining thirty percent between occupations are similar to Feenstra and Hanson's finding.

intermediates and their integration into final products. Ignoring these intersectoral linkages imposes an important limitation to the investigation of skill upgrading. This paper suggests that technology-skill complementarities across sectors can explain more skill upgrading in U.S. manufacturing than high-tech capital, R&D intensity, or outsourcing.

A tale of two sectors

To illustrate my finding, I contrast the divergent experiences of two sectors. Both began with a white-collar worker share similar to the manufacturing average (24% in 1967), but took very different paths thereafter: One revolutionizing its products, while the other turned out an unchanging artifact. The first sector, Calculating and Accounting Equipment (SIC 3578), experienced major skill upgrading, with the share of white-collar workers increasing from 23% in 1967 to 58% in 1992. In contrast, this share stagnated at 20% in the second sector, Truck Trailers (SIC 3715), lagging far behind the manufacturing average that grew to 31% in 1992.[10] Table 3.1 shows for both sectors the six most important intermediate inputs and the white-collar labor share employed in their production.

Over the period 1967-1992, the Calculating and Accounting Equipment sector underwent major innovations, above all the switch from mechanical (wiring, metal, machines) to high-tech components (computing, electronic, and semiconductors). This transition is reflected by the changing input shares a_j in Table 3.1. The producers of high-tech components, in turn, experienced skill upgrading, as reflected by changes in h_j. For example, semiconductors were produced with 32% of white-collar workers in 1967 as compared to 51% in 1992. Less skill-intensive (and less innovative) inputs like wiring devices, on the other hand, were important in 1967 but nonrelevant in 1992. Therefore, skill upgrading in the Calculating and Accounting Equipment industry went hand in hand with innovation and skill upgrading in the production of its intermediate inputs. This example provides a strong case for product innovation driving skill demand in many sectors, as opposed to the commonly studied within-sector effects of process innovation. As Pavitt (1984, p.350) puts it, referring to the same sector:

> "Innovative activities are in fact heavily concentrated on product innovation: no amount of process innovation in, for example, the production of mechanical calculators would have made them competitive with the product innovations resulting from the incorporation of the electronic chip."

Things look different in the Truck Trailer industry, where input mix and skill intensity of input production changed little. As the lower part of Table 3.1 shows, input shares are very stable over time – a truck trailer in the 1990's is not much different from one three decades earlier. Moreover, sectors supplying intermediate inputs for truck trailer production experienced minor or no skill upgrading, indicating little product innovation. Product monotony goes hand in hand with stagnation of the workforce composition: The white-collar labor share remained unchanged throughout 25 years.

Intermediate input linkages

While intermediate linkages play no role in the SBTC literature, studies concerned with linkages concentrate on labor productivity rather than skill bias and wage inequality. Although intermediate inputs

[10]The data are from the NBER Manufacturing Industry Database. See section 3.3 for details.

Table 3.1: Two sectors: Intermediate input shares and skills used in intermediate production

Calculating and Accounting Equipment (SIC 3578)

	1967		1992	
Input (j)	a_j	h_j	a_j	h_j
Electronic computing equipment	.018	.419	.364	.663
Miscellaneous electronic components	.055	.267	.232	.358
Semiconductors & related devices	.021	.322	.095	.507
Wiring devices	.116	.249	.000	.285
Office machines	.110	.284	.000	.500
Metal stampings	.038	.172	.068	.237

Truck Trailers (SIC 3715)

	1967		1992	
Input (j)	a_j	h_j	a_j	h_j
Motor vehicle parts & accessories	.237	.180	.201	.217
Aluminum rolling & drawing	.122	.208	.144	.241
Blast furnaces & steel mills	.150	.186	.115	.230
Tires & inner tubes	.065	.230	.121	.183
Fabricated rubber products	.039	.255	.053	.265
Sawmills & planing mills, general	.030	.088	.037	.146

Notes: Data from U.S. Input-Output tables and the NBER Manufacturing Industry Database. See section 3.3 for details. a_j: The respective sector's expenditure share for input j (relative to total expenditures for manufacturing inputs purchased outside the same sector); ordered by average importance in 1967/92. h_j: Share of white-collar workers in production of input j

account for more than half of all costs, the literature has not taken notice of intersectoral linkages as an explanation for skill demand. Input-output tables for the United States show that industries' expenditure share for intermediate inputs is stable over time, largest in manufacturing (57%), and smallest in services (43%). The remaining expenditures include employee compensation (about 30%) and payments to capital (about 16%).[11] Studies arguing that a capital-skill complementarity is responsible for skill upgrading therefore focus on a relatively small component of the final product's value.[12] The approach applied in this paper is strictly separated from the capital-skill complementarity literature. While the latter analyzes SBTC related to capital (or investment) goods, my analysis is based on intermediate input-output linkages that by construction do not include investment.

The importance of input-output linkages for economic development has been investigated in an ample

[11]These two, together with the minor component 'Indirect business tax and nontax liability' make up value added (on average 47% of all expenditures). All percentage values are derived from the 1992 U.S. input-output table from the Bureau of Economic Analysis. Numbers are very similar in other years.

[12]The hypothesis of capital-skill complementarity has been formalized by Griliches (1969). Krusell et al. (2000) argue that the stock of capital equipment, measured in efficiency units, is complementary to skilled labor, accounting for much of the variations in the skill premium over the last 30 years. This result has been challenged because it disappears upon the inclusion of a time trend, which is the case in my analysis, as well.

literature pioneered by Leontief (1936) and Hirschman (1958). Ciccone (2002) shows that small increasing returns at the firm level can translate into large effects on aggregate income when industrialization goes hand in hand with the adoption of intermediate-input intensive technologies. In a recent contribution, Jones (2007) analyzes this point more deeply, underlining the role of linkages and complementarities to explain large cross-country income differences. In his paper, input-output linkages give rise to a multiplier effect in production that augments productivity differentials; the latter are in turn explained by complementarities along the production chain. Multipliers have also been used to explain the growth in the trade share of output, or the cyclical behavior of aggregate productivity.[13] However, this paper is the first to investigate the role of intersectoral linkages for skill upgrading.

Innovation linkages across sectors

Linkages across industries alone need not imply connected skill requirements.[14] What makes the proposed point plausible is the (above discussed) innovation-skill complementarity *within* sectors, combined with strong empirical evidence showing that innovation is transmitted *across* sectors through input-output linkages.

Research from the 1980's provides substantial evidence for technology linkages across sectors. Scherer (1982) using U.S. patent data, and Pavitt (1984) using British innovation data, implement a methodology first proposed by Schmookler (1966). They construct what can be considered the technological equivalent of an input-output table, identifying sectoral R&D expenditures, as well as the amount of each sector's R&D that is passed to other sectors in the form of product-embedded innovation. In this context, product innovations are by definition used outside their sector of origin, and process innovations are used inside their sector. For example, in the United States 86% of all R&D expenditures in the Lumber and Wood sector improved production processes, and only 14% of innovations left this sector in the form of improved products. The opposite holds for Industrial Electrical Equipment, where 85% of R&D was devoted to product innovation, benefitting sectors that use electrical equipment. Both Scherer and Pavitt confirm the overall prevalence of product innovations, which account for 73.8 percent of total R&D outlays in the USA, and 75.3 percent in Great Britain. Therefore, the majority of innovations influence product characteristics and design outside their sector of origin.[15] How does this pattern of production and use of innovations compare with recent contributions to the SBTC literature that analyze innovation- and capital-skill complementarities solely *within* sectors? These studies assume that technology is created by R&D within a sector, or that it is capital-embodied, entering the sector through investment. For non-manufacturing sectors, where technical change comes mainly through the purchase of equipment, these assumptions are realistic. Within manufacturing, however, much of technical change is originating outside of a given sector and enters the sector through intermediate inputs. As Scherer (1982, p.227) emphasizes:

"If [a new product] is a producer good or intermediate sold externally, it serves to improve

[13] Yi (2003) shows that small decreases in tariff barriers multiply up to large trade increases when intermediates are traded several times during the production process. Basu (1995) argues that intermediate goods act as a multiplier for price stickiness, augmenting little firm-level rigidity to a large economy-wide price inflexibility.

[14] As pointed out by Jones (2007), one can have linkages without complementarity of inputs.

[15] Scherer (1982) also provides evidence that most productivity benefits are realized by R&D using, rather than product R&D-originating industries.

> output/input relationships or the quality of output in the buying industries. With a new turbojet engine product, for example, the R&D is performed in the aircraft engine industry, but the productivity effect often shows up in lower energy consumption or faster, quieter, and more reliable operation of equipment used by the quite distinct airlines industry. [...] to assume that the productivity-enhancing effect occurs solely within the R&D-performing industry [...] is more wrong than right, since three-fourths of all industrial R&D is devoted to new or improved products, as distinguished from processes."

This discussion underlines the existence of innovation spillovers from upstream suppliers to downstream final producers, via intermediate linkages. Does this channel exist in the opposite direction, too? That is, do innovative final producers demand cutting-edge intermediates? At the national level, this specific causal relationship is empirically difficult to separate from agglomeration economies, due to the proximity of production activities.[16] However, the literature on international spillovers and transfer of knowledge provides evidence that innovative downstream producers foster technical progress of their upstream suppliers. For example, Blalock and Gertler (2002) document vertical spillovers in the case of foreign investment in Indonesia: Subsidiaries of multinational enterprises provide technological knowledge to their local intermediate suppliers in order to reduce prices and increase competition in upstream markets.[17]

Adding the role of technology-skill complementarity

We currently know that innovation goes hand in hand with skilled labor. Moreover, innovative activity improves a sector's products. Some are used as intermediates in other sectors, generating spillovers along the production chain. So far, these two facts have been treated separately in the literature. Combining them yields an intersectoral technology-skill complementarity. The interactions of innovations and skills run in both directions, and across sectors, reinforcing one another. Individually and collectively, innovations in sectors related through input-output (I-O) linkages increase the relative demand for skilled labor (H/L) as summarized below:

<u>Upstream</u>		<u>Downstream</u>
Product Innovation$_j$	$\Leftrightarrow$	Product Innovation$_i$
$\updownarrow$	I-O linkages	$\updownarrow$
$(H/L)_j$		$(H/L)_i$

Closest to my contribution are the complementarity frameworks proposed by Milgrom and Roberts (1990) and Bresnahan et al. (2002), where the adoption of IT, work organization, product innovation, and skill upgrading reinforce each other within, but not across firms.

[16]Besides technological spillovers and intermediate input linkages, Marshall's (1927) three major sources of agglomeration also include labor pools. Robbins (2006) identifies knowledge spillovers that spread across U.S. states and vary in magnitude depending on distance and technologies.

[17]For a theoretical framework see Rodríguez-Clare (1996). Keller (2004) and Koo (2005) summarize the literature on international and local technology spillovers.

3.3 Data

Data on worker characteristics, wages, value of shipment, and real capital (equipment and structures) at the 4-digit SIC level are from the NBER-CES Manufacturing Industry Database. These data are collected from various years of the Annual Survey of Manufacturing (ASM), and have been widely used to investigate the determinants of the rise in U.S. wage inequality.[18] This database classifies employment in two broad categories: production and non-production workers. The former are 'workers engaged in fabricating, processing, assembling, inspecting, and other manufacturing', while the latter are 'personnel, including those engaged in supervision, installation and servicing of own product, sales, delivery, professional, technological, administrative, etc.' According to this classification, non-production workers are involved in innovative activities, the focus of this paper. As noted by Berman et al. (1994), the production/non-production classification closely mirrors the distinction between blue- and white-collar occupations from the Current Population Survey, which in turn closely reflects educational levels as high school vs. college. In the following, I refer to non-production (white-collar) workers as high-skilled labor H and to production (blue-collar) workers as low-skilled labor L.

The Bureau of Economic Analysis' (BEA) Input-Output Use Tables specify expenditures of each industry i for intermediate inputs purchased from industry j. The BEA provides U.S. input-output (I-O) data at the 4-digit SIC level in 5 year periods between 1967 and 1992. For some sectors, the level of aggregation or coverage changes over time. I account for this by aggregating sectors, and match the resulting I-O panel to the ASM's 1987 SIC classification.[19] This yields a panel of 358 manufacturing industries over the period 1967-1992. For each industry, the panel contains production and non-production employment and wages, value of shipment with the corresponding deflator (1987=1), real capital equipment and structures (all from the ASM), and the purchases of industry i from sector j (from the BEA I-O data). All figures provided in the BEA's I-O Use Tables are in nominal dollars. I use the shipment deflators provided by the ASM to calculate, for every manufacturing industry i, its real expenditures for inputs from each manufacturing industry j in year t, X_{ij}^t.[20]

Constructing the measure of input skill intensity

To construct a measure of skills embedded in intermediate inputs, I first derive intermediate input shares from the real I-O expenditure data X_{ij}^t. Let $X_i^t = \sum_{j \neq i} X_{ij}^t$ represent total expenditures for manufacturing inputs purchased by industry i outside the same industry in period t. The time-varying intermediate input shares are then given by $a_{ij}^t = X_{ij}^t / X_i^t$. These exhibit substantial fluctuations over time, mostly due to one-time outliers in the six benchmark years. For example, in 1967 'Paperboard containers and boxes' accounted for 3.4% of the manufacturing inputs in the 'Chocolate and cacao products' sector. This number more than quadrupled 5 years later (13.4%), stabilizing at 6.5% thereafter until 1992. Another example is 'Communication equipment', used in 'Guided missiles and space vehicles' production. The corresponding a_{ij} grows from 4% in 1967 to 47% in 1977, then falling back to 5% in 1992. There is no

[18] Examples include Berman et al. (1994), Autor et al. (1998), and Feenstra and Hanson (1999). See Bartelsman and Grey (1996) for a documentation of these data.

[19] For example, paper mills (SIC 2621) and paperboard mills (SIC 2631) are available separately in the I-O data until 1982, but aggregated from 1987 on. I treat these data as one sector, 'paper and paperboard mills' over the full sample period. Detailed sector correspondences are available upon request.

[20] Bartelsman and Grey (1996) use the same method to derive real material (or input) costs.

reason to believe that these numbers reflect physical input shares. The paper wrappings around chocolate did not become thicker in 1972. Measurement error as well as fluctuations in relative input prices, imperfectly corrected by the deflators, appear to be reasonable explanations.[21] Therefore, I use average real input shares $\bar{a}_{ij} = \sum_{t=67}^{92} a_{ij}^t$ between 1967 and 1992 as a baseline. This approach can be interpreted to reflect an underlying Cobb-Douglas technology that keeps expenditure shares constant over time (or a Leontief that has the same effect under stable relative prices).[22] Input skill intensity is then defined as

$$\sigma_i^t = \sum_{j \neq i} \bar{a}_{ij} h_j^t \qquad (3.1)$$

where $h_j^t \equiv H_j^t/(H_j^t + L_j^t)$ denotes the share of white-collar workers employed in the production of input j.[23] I exclude inputs purchased within the same sector ($j = i$) for two reasons. First, this avoids that skilled workers employed in sector i itself enter its measure of input-embedded skills σ_i, which would bias my results. Second, I am concentrating on product innovation entering a sector via intermediates purchased from outside, rather than process innovation generated within a sector.

A potential concern arises because inputs X_{ij} (and thus input shares a_{ij}) contain imports from abroad, while the corresponding skill shares h_j are measured in U.S. sectors.[24] However, the resulting measurement error of σ_i is likely to be minor. The share of imports in non-energy intermediates during my sample period is relatively small, growing from 4% in 1967 to 13% in 1992 (see Appendix A1). Moreover, most U.S. imports of intermediates in this period were sourced from other OECD countries with similar skill intensities. Finally, having a noisy measure of input skill intensity creates attenuation bias against finding skill complementarities across sectors.

By construction $\sigma_i \in [0, 1]$ is the weighted average share of non-production workers involved in the production of sector i's intermediate manufacturing inputs. An alternative measure of input skill intensity is obtained by excluding those inputs that are purchased from the same two-digit SIC industry as the good being produced. I implement this idea by restricting the four-digit industry subscripts i and j in (3.1) to be outside the same two-digit SIC industry. This measure addresses the concern that skill upgrading may happen simultaneously in similar industries, which would imply a correlation of input and final production skill intensities when similar sectors buy each other's inputs. The resulting measure is labeled σ_i^{2d}.

Table 3.2 provides a list of the twenty industries with the smallest and the largest increase in input skill

[21] Out of the about 128,000 $i \times j$ input shares, 7,000 are nonzero throughout the sample period. Their average coefficient of variation over the six sample benchmark years is 0.67. Less than 1/3 have a time trend that is significant at the 10% level.

[22] This would be a strong assumption if input shares shifted systematically towards more (or less) skill intensive industries, which is not the case. In section 3.4.2 I use the time-varying a_{ij} to decompose input skill intensity into input-mix and skill-mix components. This analysis shows that practically all the increase in input skill intensity between 1967 and 1992 is due to skill upgrading in input production at constant input shares (skill mix), rather than changing input shares (input mix). Section 3.4.2 also provides an extended empirical analysis with time-varying input shares, showing that the ITSC is robust to this specification.

[23] Alternatively, σ_i^w can be calculated, using wage-bill instead of employment shares: $h_j^w \equiv w_{H,j} H_j/(w_{H,j} H_j + w_{L,j} L_j)$, where $w_{H,j}$ and $w_{L,j}$ denote white- and blue-collar wages, respectively. Regression results change only very little when using σ_i^w.

[24] Unfortunately, the BEA provides import matrices only from 1997 on. But even these numbers are approximations and do not include the source country. Actual data on the domestic/imported content of an industry's intermediate inputs are, for the most part, not available. Import matrix estimates are typically based on the assumption that the share of imports in total domestic consumption of a commodity applies to each industry that uses the commodity (proportionality assumption).

intensity σ_i^{2d} for the period 1967-92.[25] The reported ranking seems sensible. The industries with smallest changes (or declines) in input skill intensity are mainly textile and food industries. These tend to use primary inputs, which in turn changed little or dropped in terms of white-collar employment shares. Most industries that experienced the largest increase in inputs skill intensity also appear sensible. These include various electronic, computing, and communication equipment, as well as aircraft and space industries, all of which intensively use high-tech inputs that experienced innovation and skill-upgrading throughout the last decades.[26]

Table 3.2: The twenty industries with smallest and largest change in input skill intensity

Smallest change in σ_i^{2d} 1967-92		Largest change in σ_i^{2d} 1967-92	
$\triangle\sigma_i^{2d}$	Industry description	$\triangle\sigma_i^{2d}$	Industry description
-.045	Leather tanning & finishing	.074	Carbon black
-.023	Tire cord & fabrics	.074	Ceramic wall & floor tile
-.022	Yarn mills & finishing of textiles	.075	Watches, clocks, & parts
-.011	Women's hosiery, except socks	.075	Photographic equipment & supplies
.001	Carpets & rugs	.076	Paperboard containers & boxes
.006	Cordage & twine	.076	Primary aluminum
.009	Thread mills	.076	Primary nonferrous metals
.010	Knit fabric mills	.077	Ordnance & accessories
.020	Hosiery	.079	Steel pipe & tubes
.021	Manufactured ice	.079	Search & navigation equipment
.021	Footwear cut stock	.080	Aircraft
.025	Leather gloves & mittens	.081	Wood preserving
.026	Knitting mills	.082	Paper & paperboard mills
.027	Schiffli machine embroideries	.082	Calculating & accounting equipment
.027	Malt beverages	.086	Typesetting
.028	Truck trailers	.089	Instruments to measure electricity
.028	Mobile homes	.090	Pulp mills
.028	Bottled & canned soft drinks	.091	Electronic computing equip.
.029	Frozen fruits & vegetables	.093	Guided missiles & space vehicles
.029	Fertilizers, mixing only	.111	Electromedical equipment

Note: Reported input skill intensities are rounded from seven digits to three digits.

[25]The sectoral *levels* of input skill intensity are not important for my empirical results – they are taken up by fixed effects in the regressions. Thus, I report changes rather than levels. The ranking is similar when based on the measure σ_i.

[26]Pulp Mills and Paper & Paperboard Mills do not seem to fit this pattern. Part of the increase in input skill intensity is explained by their dependence on Industrial Chemicals (about 1/4 of all inputs), which experienced skill upgrading from 35 to 44 percent. However, another part is due to accounting, rather than real skill upgrading in input production. Both industries depend heavily on inputs from Logging, with the corresponding input shares 36 and 28 percent, respectively. The non-production labor share in Logging rose from 4.4% to 17.0%. A possible explanation for this increase is offered by the Occupational Employment Statistics from the Bureau of Labor Statistics, which provides detailed occupation data from 1999 on. According to these data Logging involved about 22 percent of employment in transportation activities in 1999. The ASM classifies transportation as non-production labor. The rising importance of transportation is relative rather than absolute, because overall employment in Logging fell. Because few sectors depend on inputs from Logging, this problem is an isolated one. Moreover, my results are robust to splitting the sample into sectors with falling and increasing overall employment.

The measure of input differentiation

In order to identify the degree of differentiation for each sector's inputs, I use data from Rauch (1999). Rauch groups goods into 1,189 industries according to the 4-digit SITC Rev. 2 system. An industry's product is classified as being differentiated if it is neither traded on an organized exchange nor reference priced in trade publications.[27] I aggregate the Rauch data into the 358 SIC industries of my sample. This procedure yields data on the fraction of each industry's output that is differentiated.[28] Using this information, along with the input shares derived above, I define the degree of input differentiation:

$$\kappa_i = \sum_{j \neq i} \bar{a}_{ij} \; R_j^{\text{diff}} \tag{3.2}$$

where R_j^{diff} is the proportion of input j that is classified as differentiated. The measure κ_i is therefore the weighted average share of a sector's inputs (purchased outside the same sector) that are differentiated. This variable is similar to Nunn's (2007) measure of relationship specificity; but Nunn uses Rauch's classification in a different context, showing that countries with good contract enforcement specialize in the production of goods that require relationship-specific investments.

Data on product innovation

I use data from Scherer (1982) to derive, for each industry, its share of R&D spent for product innovation, π_i^{prod}.[29] In the empirical analysis π_i^{prod} serves to investigate the relationship between product innovation and product differentiation, given by R_i^{diff} described above. In order to perform this analysis, I match my 4-digit SIC code to Scherer's 36 manufacturing industries and aggregate R_i^{diff} to this level of detail, using sectoral shipments as weights. The resulting sample includes π_i^{prod} and R_i^{diff} for 34 manufacturing sectors (2 observations of π_i^{prod} are missing). π_i^{prod} has mean .66 and standard deviation .27. The share of product innovation is smallest in primary industries like wood products, ferrous metals, or petroleum, and largest in various machinery and equipment industries, including photo, medical instruments, communication and construction equipment.

Additional control variables

In the empirical analysis I include several variables that have been previously used to explain increasing wage inequality. In the following I describe these variables briefly. Appendix A.1 provides more detail. Krusell et al. (2000) argue that the stock of capital equipment is complementary to skilled labor. To control for this capital-skill complementarity, I include real equipment and real structures per worker, k^{equip} and k^{struct}, respectively. Data on research and development (R&D) intensity are from the National Science Foundation (NSF). Following Autor et al. (1998), I use lagged R&D intensity ($R\&D_{lag}$) in the

[27] Rauch provides liberal and conservative estimates. I use the former, but none of the results presented in the following depend on this choice.

[28] Nunn (2007) describes the construction of a crosswalk from the 4-digit SITC to the BEA's 1987 4-digit SIC classification. He kindly shared his data. These aggregate into 302 sectors of my sample. For the remaining 56 sectors I use a correspondence from 4-digit SITC to 4-digit SIC provided by Pamela Lowry (downloadable from Jon Haveman's Industry Trade page). Like Nunn, I apply equal weights when aggregating SITC industries to the SIC classification.

[29] Appendix A.1 explains the corresponding methodology in detail.

regressions.[30] I use data from the BEA to construct sectoral shares of high-technology capital (*HT/K*) and office, computing & accounting equipment (*OCAM/K*).[31] Feenstra and Hanson (1999) document a substantial impact of foreign outsourcing of intermediate inputs on relative wages. I calculate their broad (OS^{broad}) and narrow (OS^{narr}) measures of outsourcing for the years and sectors included in my sample. Feenstra and Hanson argue that the narrow measure – from within the same two-digit industry – best captures the idea of outsourcing. For example, the import of steel by a U.S. automobile producer is normally not considered as outsourcing, while it is common to think of imported automobile parts by that company as outsourcing. Following this reasoning, I use OS^{narr} in most regressions, including OS^{broad} in the robustness checks.

Table 3.3 reports the pairwise correlations between two measures of input skill intensity (σ_i and σ_i^{2d}) and the most prominent control variables. As in most of the following analyses, these correlations are obtained after controlling for industry and time fixed effects. The two measures of input skill intensity are highly correlated with one another, and are also correlated with control variables commonly used in the SBTC literature. Industries using skill-intensive intermediates tend to be capital and R&D intensive, employ high-tech capital, and outsource the production of their intermediates.

Table 3.3: Correlations between input skill intensity and control variables

	Input skill intensity		Capital per worker		High-Tech	R&D/ sales	Out-sourcing
Measure	σ_i	σ_i^{2d}	k^{equip}	k^{struct}	HT/K	$R\&D_{lag}$	OS^{narr}
σ_i	1						
σ_i^{2d}	.66***	1					
k^{equip}	.12***	.15***	1				
k^{struct}	.05**	.07***	.70***	1			
HT/K	.20***	.14***	-.03	-.01	1		
$R\&D_{lag}$	.18***	.13***	-.01	.03	.39***	1	
OS^{narr}	.13***	.11***	.05**	.04*	.03	.10***	1

Notes: Reported numbers are pairwise correlation coefficients, controlling for sector and time fixed effects.
Key: *** significant at 1%; ** 5%; * 10%.

3.4 Empirical Results

Complementarity implies correlation. It is irrelevant for the ITSC whether we think that "downstream skills and innovation cause upstream innovation and skills" or the other way around. If new technology and skills are complements along the production chain, skill upgrading at one level affects innovation

[30]The first (lagged) period is 1963, implying a 4-year lag. All other lags are 5 years. Because industrial R&D intensity tends to be persistent over time, working with lagged or contemporaneous R&D makes very little difference to the nature of the results.

[31]Both technology measures are widely used in studies of wage inequality. See, in particular, Autor et al. (1998) and Feenstra and Hanson (1999). The capital stock data are likely measured with substantial error, and are often not measured directly but inferred from employment data, assuming relationships between occupations and capital-type usage. See Becker et al. (2006) for a discussion. This implies an upward bias of computer capital's impact on skill upgrading, stacking the odds against finding an important contribution of input skill intensity.

and skill demand in both directions, upstream and downstream. Thus, the complementarity theory can be investigated in either causal direction. I follow the common identification strategy in the SBTC literature and use the high-skill labor share, h_i, as dependent variable.

First I show that the novel fact presented in the introduction is not an artifact: the correlation between input skill intensity and final production skills is robust to a variety of additional controls and specifications. After this, I provide evidence suggesting that the ITSC works through product innovation. Finally, I examine the ITSC's importance for skill upgrading and address endogeneity issues.

3.4.1 Correlation of Skill Intensities across Sectors

A first look at the data was provided above by Figure 3.1, plotting a cross-section of h_i against σ_i, where both variables are calculated in 1992. The corresponding regression, including a constant term, yields a highly significant coefficient: $\beta = .957$, with a (robust) standard error of .101. Two concerns arise. First, the observed correlation may be due to unobserved sectoral characteristics that drive both skill demand and input skill intensity.[32] Second, when using a panel, the correlation between h_{it} and σ_{it} may be spurious, driven by a general trend of skill upgrading. To address these concerns, I now turn to using the full panel, controlling for time and sectoral fixed effects. In addition, I control for variables that have been previously identified as influencing h_i. I estimate the following equation

$$h_{it} = \alpha_i + \alpha_t + \beta\sigma_{it} + \gamma\mathbb{Z}_{it} + \varepsilon_{it} \tag{3.3}$$

where α_i and α_t denote industry and time fixed effects, respectively; $\mathbb{Z}_{it}$ are control variables, and ε_{it} denotes the error term, capturing measurement error and unobserved drivers of the skilled labor share. The first column of Table 3.4 shows regression (3.3) with sectoral and time fixed effects. The coefficient on σ_{it} is highly significant. The number of observations represents the full sample of 6 years $\times$ 358 sectors = 2148. I report two frequently used measures for the goodness of fit: One including the variation explained by sectoral fixed effects (R^2), and the other assessing the model's fit after accounting for sectoral dummies (R^2 within). The former is close to one, while the latter implies that the regressions presented in Table 3.4 account for roughly half of the variation of h_{it} within sectors over time.

Next, I control for capital endowments as determinants of skill upgrading. Krusell et al. (2000) find a strong positive impact of capital equipment on skill demand for the aggregate U.S. economy. As column 2 shows, this finding is not reproduced at the detailed industry level; the coefficient on k^{equip} has the wrong sign and is significant at the 10% level.[33] I also include capital structures, which are skill-neutral in the setup of Krusell et al. (2000), and on the verge of influencing skill upgrading significantly in my sample. The share of high-tech capital correlates significantly positively with the proportion of skilled labor, resembling the well-documented complementarity. This variable has more explanatory power than

[32] One such story would be that both skill-intensity of sectors and the inputs they use are 'naturally given' (e.g., determined by technological history) and independent of innovative activity. Suppose that 'ancient' sectors are low-skill intensive, buying mainly 'ancient' inputs, while 'modern' sectors employ skilled workers and purchase 'modern' inputs. This would yield the observed correlation in the absence of intersectoral technology-skill complementarities.

[33] This supports the critical view of Krusell et al.'s results, which disappears when a time trend is included. In fact, if I only include k^{equip} and sectoral dummies as explanatory variables in (3.3), the coefficient on k^{equip} is positive and highly significant. As soon as other controls or time dummies are included, the coefficient becomes insignificant. Note, however, that sector-specific real equipment data from the ASM used in my sample do not include the quality-adjustment that Krusell et al. apply at the aggregate level.

Table 3.4: Final production skills, input skill intensity, and controls. Dependent variable is h_{it}.

Input skill measure	Baseline: σ_i				σ_i^{2d}	σ_i^w
	(1)	(2)	(3)	(4)	(5)	(6)
Input skill intensity: σ_i	.834***	.658***	.558***	.665***	.502***	.548***
	(.156)	(.145)	(.126)	(.066)	(.170)	(.148)
Structures per worker: k^{struct}		.259*	.191	.249**	.232*	.175
		(.134)	(.117)	(.120)	(.122)	(.118)
Equipment per worker: k^{equip}		-.114*	-.0992	-.168**	-.103	-.0687
		(.067)	(.062)	(.070)	(.064)	(.061)
High-Tech capital: HT/K		.716***	.600***	.410***	.618***	.614***
		(.134)	(.151)	(.125)	(.150)	(.153)
Office equipment: $OCAM/K$		-.0692	.0102	.0576	.0371	.016
		(.294)	(.316)	(.324)	(.308)	(.316)
R&D intensity $R\&D_{lag}$			.401**	.322*	.461***	.363**
			(.163)	(.193)	(.158)	(.163)
Outsourcing: OS^{narr}			.146***	.122**	.161***	.139***
			(.050)	(.051)	(.053)	(.048)
Sector fixed effects	yes	yes	yes	yes	yes	yes
Time fixed effects	yes	yes	yes	no	yes	yes
R^2	.97	.98	.98	.98	.98	.98
R^2 (within)	.50	.55	.57	.56	.56	.57
Observations	2148	2148	2089	2089	2083	2089

Notes: Clustered standard errors (by sector) in parentheses. Key: *** significant at 1%; ** 5%; * 10%. All regressions are weighted by sectors' average share in total manufacturing employment 1967-92.

the alternative measure that only includes office, computing, and accounting equipment. Finally, and most important for my results, the coefficient of input-skill intensity is robust to the inclusion of capital controls. The same holds when further controls are included, as shown in column 3. The sample size is now 2089 due to missing observations in the outsourcing measure. Both lagged R&D intensity and outsourcing have a significantly positive correlation with skilled labor in final production, which confirms previous findings (Machin and van Reenen 1998, Feenstra and Hanson 1999). Column 4 shows the results without time dummies. As expected, because of the general skill upgrading over time, the coefficient of input skill intensity is now slightly bigger. Remarkably, the coefficient of capital equipment is significantly negative, providing further evidence against a capital-skill complementarity.

The last two columns of Table 3.4 present regression results for alternative measures of input skill intensity, including all controls. In column 4, σ_i^{2d} is used, excluding inputs purchased within the same 2-digit industries. This specification addresses the concern that common trends or technology shocks may drive skill upgrading in similar industries, biasing β upwards when these industries are linked via input-output relationships. The more conservative measure comes along with a cost: σ_i^{2d} discards a substantial part of intersectoral linkages, since sectors purchase on average 35% of their inputs within the same 2-digit category.[34] Therefore, σ_i^{2d} is a more noisy measure of input skill intensity and likely subject to attenuation bias. However, the coefficient β is only slightly smaller than in the previous specifications and

[34] One sector, 'Special product sawmills' (SIC 2429) purchases all inputs within the same 2-digit category. The corresponding σ_i^{2d} is therefore missing in all 6 benchmark years, leaving 2083 observations.

still highly significant. I implement two additional ways to address the common-shock concern. Both are based on specification 3 and are not reported in the table. First, I use the 5-year lag of σ_i. The coefficient on $\sigma_{i,t-5}$ is highly significant, .366 (.095), with all other coefficients very similar to those reported in columns 3 and 5. This finding mitigates the common-shock concern – to maintain it, one would have to argue that downstream skill demand reacts half a decade later than its upstream counterpart to the same shock. Second, I include time dummies at the 2-digit industry level. These absorb any industry-specific shocks to skill demand, such that the coefficient β only reflects the variation of detailed 4-digit sectors relative to the corresponding 2-digit industries. Even with this restriction, the coefficient remains highly significant and of similar magnitude, $\beta = .401$ (.185). Finally, column 6 uses σ_i^w, where skills in input production are measured with the wage-bill, instead of the labor share of skilled workers (see footnote 23). Berman et al. (1994) propose the wage-bill share as an alternative measure of skill demand, because it also captures skill upgrading within either category – production or non-production workers. The results obtained with σ_i^w are very similar to the ones with σ_i.

3.4.2 Robustness of the Correlation

The robustness of my results to alternative measures of input skill intensity, σ_i, σ_i^{2d}, and σ_i^w has been verified in Table 3.4. These measures were all calculated based on constant input shares, i.e., stable linkages over time. In this section, I first show that my results are robust to including input skill intensity measures based on changing input shares. Second, I test the sensitivity and robustness of my estimates to alternative specifications.

Input skill intensity with changing input shares

Because input shares a_{ij} vary substantially over time, mostly due to one-time outliers, my baseline input skill intensity measures are derived based on average input shares $\bar{a}_{ij}$. Now, I use the time-varying a_{ij} to construct the input skill intensity measure $S_i^t = \sum_{j \neq i} a_{ij}^t h_j^t$. This variable can be decomposed into three parts. First, a skill component σ_i^t, as defined in (3.1), representing constant input expenditure shares with changing skilled labor shares of suppliers. Second, an input-mix component $\tau_i^t = \sum_{j \neq i} a_{ij}^t \bar{h}_j$, reflecting varying input shares with constant skilled labor shares of suppliers. This variable grows over time if sector i switches its input mix towards more skill intensive intermediates. Finally, a covariance component $\tau_i^t = \sum_{j \neq i} (a_{ij}^t - \bar{a}_{ij})(h_j^t - \bar{h}_j) - \sum_{j \neq i} \bar{a}_{ij} \bar{h}_j$, which grows if sector i switches its input mix towards sectors whose skill intensity rises over time.[35] Note that $S_i^t = \sigma_i^t + \tau_i^t + \rho_i^t$. The skill component σ_i is by far the most important contributor to increases in S_i^t between 1967 and 1992. The weighted average of S_i^t increases from 21.2 to 27.6 percent. Of this 6.4% rise, 6.2% are due to σ_i, 1.3% to τ_i, and -1.1% to ρ_i. As Table 3.5 shows, the coefficient of σ_i does not change when the two additional variables are used – it is still above 0.5.

Once the usual controls are included, neither τ_i nor ρ_i are significant, as shown in the second and third column of Table 3.5. This result was to be expected, given the noise in the input shares used to calculate these variables.[36] Similarly, we expect attenuation bias and therefore a smaller coefficient when using

[35]The term $\sum_{j \neq i} \bar{a}_{ij} \bar{h}_j$ is a constant for each sector i and does not influence estimation results in the presence of sectoral fixed effects.

[36]Less than 1/3 of all input shares have a time-trend that is significant at the 10% level. In an additional check not presented

Table 3.5: Input skill intensity with time-varying input shares. Dependent variable is h_{it}.

Input skill measure	σ_i		σ_i^{2d}	S_i	
	(1)	(2)	(3)	(4)	(5)
Input skill intensity					
Skill component: σ_i / σ_i^{2d}	.832***	.562***	.511***		
	(.150)	(.124)	(.161)		
Input mix component: τ_i / τ_i^{2d}	.110	.068	-.011		
	(.077)	(.061)	(.079)		
Covariance component: ρ_i / ρ_i^{2d}	.725*	.236	.224		
	(.389)	(.394)	(.466)		
All together: $S_i = \sigma_i + \tau_i + \rho_i$				.189***	.325***
				(.059)	(.049)
Controls	no	yes	yes	yes	yes
Sector fixed effects	yes	yes	yes	yes	yes
Time fixed effects	yes	yes	yes	yes	no
R^2	.97	.98	.98	.97	.97
R^2 (within)	.51	.57	.56	.56	.53
Observations	2090	2089	2083	2089	2089

Notes: Clustered standard errors (by sector) in parentheses. Key: *** significant at 1%; ** 5%; * 10%. All regressions are weighted by sectors' average share in total manufacturing employment 1967-92. Controls include the following variables: Structures per worker (k^{struct}), Equipment per worker (k^{equip}), High-Tech capital (HT/K), Office and computer capital ($OCAM/K$), R&D intensity ($R\&D_{lag}$), and Outsourcing (OS^{narr}).

the composite skill intensity S_i. Columns 4 and 5 show this result with and without time dummies. The coefficients on S_i are, however, still highly significant.

Alternative specifications and further controls

Alternative specifications comprise running the regression in changes, including further controls, and restricting the sample to single years, analyzing cross-sections rather than a panel. Table 3.6 presents the results. Therein, I include the computer capital share $OCAM/K$ and the *difference* between high-tech and computer capital share ($HT/K - OCAM/K$), which represents the fraction of capital services derived from various high-technology assets other than office, computing and accounting machinery. Feenstra and Hanson (1999) suggest this specification, and a similar one for outsourcing: the difference between the broad and narrow measures $OS^{\text{broad}} - OS^{\text{narr}}$, representing the intermediate inputs from outside the two-digit purchasing industry that are sourced from abroad.

The first column of Table 3.6 runs the baseline regression in changes, instead of including fixed effects. All variables are in 5-year differences.[37] The corresponding coefficient on input skill intensity is very similar to the one obtained above, and again highly significant. In column 2, I turn back to estimating levels, including fixed effects and all previously used controls. Additionally, I control for various other variables that potentially drive skill demand. First, broad outsourcing (as difference to narrow). Second, two measures of the 'complexity' of production processes: the variety of inputs used in

here, I calculate τ_i and ρ_i using changing input shares when the time-trend is significant, and average shares otherwise. Under this method, τ_i is significant at the 5% level when all controls are included, while the coefficient of σ_i remains unchanged.

[37] R&D intensity is also calculated in actual differences, rather than differences of the lagged variable.

Table 3.6: Robustness analysis. Dependent variable is h_i.

Input skill measure	σ_i		σ_i^w	σ_i^{2d}	
	(1)	(2)	(3)‡	(4)	(5)
	Changes	Additional Controls	Wage bill	1967 only	1992 only
Input skill intensity: σ_i / σ_i^{2d} / σ_i^w	.621***	.468***	.574***	.562*	.467***
	(.062)	(.123)	(.201)	(.324)	(.139)
Structures per worker: k^{struct}	.255*	.291**	.353***	- .557	.941***
	(.136)	(.121)	(.135)	(1.317)	(.337)
Equipment per worker: k^{equip}	-.0872	-.107*	-.266***	.580	-.383**
	(.090)	(.064)	(.090)	(1.442)	(.176)
Office equipment: $OCAM/K$	.107	.588	.481	3.884***	4.637***
	(.165)	(.380)	(.492)	(1.306)	(.460)
High-Tech capital: difference $(HT/K - OCAM/K)$	.124	.579***	.490***	4.238***	2.011***
	(.109)	(.152)	(.186)	(.898)	(.497)
R&D intensity $R\&D_{lag}$	.323**	.268	.468***	1.077	.144
	(.161)	(.165)	(.181)	(1.027)	(.364)
Outsourcing: OS^{narr} (narrow)	.0651*	.159***	.157**	-.650**	.0768
	(.039)	(.045)	(.069)	(.313)	(.100)
Outsourcing (broad): difference $(OS^{\text{broad}} - OS^{\text{narr}})$		.0944*		-.0369	.0781
		(.055)		(.570)	(.129)
Many inputs: $I_i^{n_i > \bar{n}}$		.00177		.00892	.0352***
		(.005)		(.026)	(.012)
Input variety: $(1 - H_i)$		-.00155		-.0837	.0385
		(.014)		(.107)	(.042)
Relative wage: $\ln(w_{H,i}/w_{L,i})$		-.0493***			
		(.017)			
Real shipments: $\ln(Y_i)$		.00912**			
		(.004)			
Value added share		.0330*			
		(.017)			
Sector fixed effects	no	yes	yes	no	no
Time fixed effects	no	yes	yes	no	no
R^2	.17	.98	.97	.24	.71
R^2 (within)	-	.59	.55	-	-
Observations	1731	2089	2089	328	356

‡ The dependent variable in (3) is the non-production wage bill share: $h_i^w \equiv w_{H,i}H_i/(w_{H,i}H_i + w_{L,i}L_i)$.
Notes: Robust standard errors in parentheses (for (1) - (3) clustered by sector). Key: *** significant at 1%; ** 5%; * 10%. Regressions (1) and (2) are weighted by sectors' average share in total manufacturing employment 1967-92; (3) by the average share in total manufacturing wage bill 1967-92; (4) and (5) by the sector's employment in 1967 and 1992, respectively. All variables in (1) represent 5-year differences (in this case, R&D intensity is $R\&D_t - R\&D_{t-5}$.), while levels are used in the remaining regressions.

production, measured as one minus the Herfindahl index of input concentration for each industry $(1-H_{it})$. This variable is used as a measure of a good's 'complexity' by Blanchard and Kremer (1997) to explain the decline of output when bargaining breaks down along the production chain. The other measure for production 'complexity' is an indicator function for the number of inputs, proposed by Nunn (2007). $I_{it}^{n_{it}>\bar{n}_t}$ equals one if the number of inputs n_{it} used in industry i in year t is greater than the median number of inputs used in all industries, $\bar{n}_t$. I derive both measures from the year-specific I-O tables. Since more 'complex' production processes require more coordination, I expect these variables to have a positive impact on demand for skilled labor. Third, I include the sector-specific skill premium, or relative wage, to capture differences in cost and quality of skilled workers across sectors.[38] Fourth, I control for productivity by including the real value of shipments, $\ln(Y)$.[39] This variable addresses the concern that productivity increases may be the driver of skill upgrading in final, as well as input production. Finally, the share of value added in total cost (derived from the BEA I-O data) controls for the overall importance of labor and capital (as opposed to intermediate inputs) in production. Service-oriented sectors generally have a larger value added share, and also a higher proportion of white-collar labor.

The inclusion of further control variables shown in column 2 of Table 3.6 changes neither the size nor the high statistical significance of the coefficient on input skill intensity. The last three additional controls are significant and have the expected sign. Interestingly, the positive and significant coefficient of real shipments, $\ln(Y_i)$, confirms Epifani and Gancia's (2006) hypothesis that the scale of production may be skill-biased. On the other hand, neither measure for production 'complexity' has a significant impact on skill demand.[40] The additional outsourcing measure has the expected positive sign and is significant at the 10% level. Column 3 presents the regression with the non-production wage-bill share as dependent variable. This measure is frequently used as an alternative to the purely labor based measure, as it also captures skill upgrading *within* either occupational category (Berman et al. 1994). The wage-bill regression confirms magnitude and significance of the ITSC effect.

In all panel regressions presented so far, I address the concern of inconsistent standard errors due to serially correlated observations by accounting for correlation within sectors across time (i.e., by clustering standard errors). Bertrand, Duflo and Mullainathan (2004) argue that this correction alone may not fully solve the problem and suggest collapsing the time series information into single periods as a further correction.[41] The last two columns of Table 3.6 implement this additional consistency check, presenting cross-sectional regressions for the first and the last benchmark year of the sample, 1967 and 1992. Fixed effects cannot be used in this specification, raising the concern that unobserved characteristics, like similarity of sectors, drive the correlation between input skill intensity and the skilled labor share in final production . To address this concern, I use $\sigma_i^{2\mathrm{d}}$ as the input skill intensity measure, excluding linkages

[38] Because of its endogeneity with skill demand, this variable is usually not included in regressions where the dependent variable is the share of skilled workers. Here, I merely use it as a control for possible cross-sectoral variations in the cost of labor classified as 'non-production' in the ASM. For example, a sector employing 30% delivery and sales personnel likely faces different non-production labor costs than one with 30% engineers.

[39] Feenstra and Hanson (1999) use this control variable. Results are very similar when using the natural logarithm of value added, as in Bresnahan et al. (2002).

[40] The two complexity measures vary little over time, such that the inclusion of sector fixed effects eliminates much of their variation. In fact, when running the same regression without sector dummies, the effect of $I_i^{n_i>\bar{n}}$ is positive and significant.

[41] Long time series (15 periods and more) are a major contributing factor to Bertrand et al.'s concern. Since my panel involves only 6 periods, the concern is likely of minor importance, given that I am already controlling for serial correlation.

within 2-digit industries. The corresponding coefficient is of the same magnitude as observed before, significant in 1967, and highly significant in 1992. Most control variables also confirm the previous findings. Capital equipment turns out negatively significant in the 1992 cross section.[42] In the panel, k^{equip} shows up negative and significant in some specifications. These findings together argue strongly against a broad equipment-skill complementarity. The more narrow high-tech capital variable, however, shows up significantly positive in the cross-section, as well. Finally, production 'complexity', measured by $I_i^{n_i>\bar{n}}$, has a significantly positive impact on skill demand in the 1992 cross-section.

3.4.3 Investigating the Channel of the ITSC

In this section, I investigate the hypothesis that the ITSC works through product innovation. I follow a three-step process. First, I show that sectors producing differentiated products spend relatively more R&D for product innovation, while producers of homogenous goods concentrate on innovating their own processes. This suggests that differentiated products embody more innovation than homogenous ones. Therefore, sectors using differentiated products as intermediates purchase relatively more embodied product innovation, which leads to the second step: If the ITSC works through product innovation, we expect it to be stronger for sectors that use relatively more differentiated inputs. Finally, I turn to the relationship between skills and productivity – the outcome of innovation. I show that innovative intermediates, measured by their skill content, raise productivity only if they meet skilled workers knowing to handle them. Consequently, skills in intermediate and final production complement each other in fostering innovation and productivity.

Product innovation and product differentiation

As described in section 3.3, I derive sectoral shares of R&D expenditures used for product innovation, π_i^{prod}, from Scherer's (1982) data, and match them to Rauch's (1999) data on product differentiation. This gives π_i^{prod} together with the share of products classified as differentiated, R_i^{diff}, for 34 manufacturing industries. The median of R_i^{diff} in this sample is .84. The 17 industries turning out goods with below-median product differentiation spend on average 53% of R&D for inventing new products (as opposed to processes), while this number is 80% for producers of above-median differentiated goods. After this preliminary observation, I turn to the simple regression $\pi_i^{\text{prod}} = \delta_0 + \delta_1 R_i^{\text{diff}} + \varepsilon_i$, where the last variable represents an error term. The corresponding estimate is positive and highly significant: $\delta_1 = .416$ with a robust standard error of .127 and R^2 of 0.27.[43] These findings suggest that differentiated products are more susceptive to product innovation, such that they are more readily reshaped by the innovative minds of skilled workers.

Input differentiation and ITSC

When skilled workers improve their products, the innovation passes through intermediate linkages to other sectors, where it also drives innovation and skill demand. As we have seen, purchasers of differenti-

[42] This finding is robust and also appears when only capital structures and equipment are included in the regression.

[43] The result is practically identical when using Rauch's (1999) conservative estimate to construct R_i^{diff}. Outliers are not an issue, and even excluding the 9 sectors that produce only differentiated products ($R_i^{\text{diff}} = 1$) leaves the remaining ones with a significantly positive δ_1.

ated inputs buy on average more innovation incorporated in their intermediates than users of homogenous ones. Consequently, we expect a stronger ITSC when input-output linkages involve more differentiated intermediates. The corresponding measure κ_i gives the weighted average degree of input differentiation, as described in section 3.3. To obtain a first look at the data, I use this measure to split the sample into sectors with below- and above-median input differentiation. Then I estimate regression (3.3) for the two subsamples and report the results in Figure 3.2 in the form of partial scatter plots. The vertical axis shows the variation in the skilled labor share h_i to be explained by input skill intensity σ_i, after controlling for fixed effects and statistically significant control variables (all controls that were significant in at least one specification in Table 3.4).

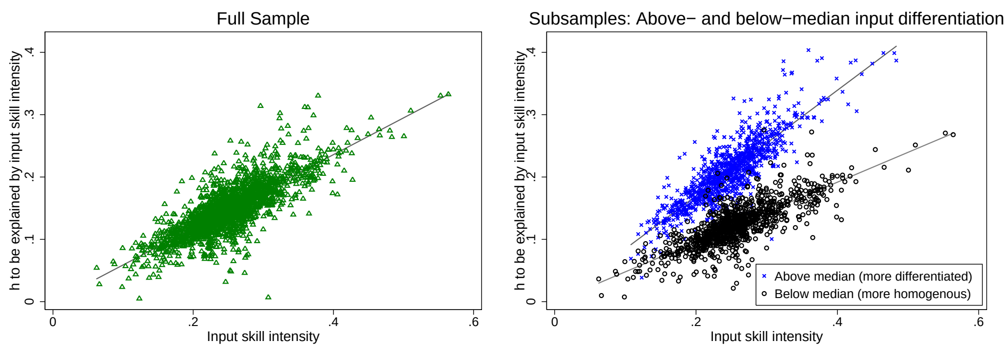


Figure 3.2: Partial scatter plots: Skilled labor share (h_{it}) vs. input skill intensity (σ_{it})

Notes: The measure of input differentiation is calculated as in (3.2), yielding a median of .52. The vertical axis shows $h_{it} - (\hat{\alpha}_i + \hat{\alpha}_t + \hat{\gamma}\mathbb{Z}_{it})$; notice that $\hat{\beta}\sigma_{it}$ does not appear in this equation. In the left panel, coefficient estimates $\hat{\alpha}_i$, $\hat{\alpha}_t$, and $\hat{\gamma}$ are obtained by estimating (3.3) for the full sample (2089 obs.), with the controls $\mathbb{Z}_{it}$ comprising k^{struct}, k^{equip}, HT/K, $R\&D_{lag}$, and OS^{narr}. In the right panel, the same methodology is applied for the two subsamples including sectors with above-median input differentiation (1040 obs.) and below-median input differentiation (1049 obs.).

The left panel of Figure 3.2 shows the partial scatterplot for the full sample, where the corresponding coefficient from regression (3.3) is $\hat{\beta} = .590$.[44] This plot also shows that the positive correlation between input skill intensity and final production skills is a broad phenomenon, not driven by outliers. The right panel repeats the exercise for two subsamples, one with sectors purchasing relatively homogenous inputs (below-median κ_i) and the other comprising sectors that use more differentiated inputs (above-median κ_i). These first results are in favor of the hypothesis that the ITSC is stronger for sectors using more differentiated inputs than for those using more homogenous ones; the corresponding coefficients are $\hat{\beta}^{\text{diff}} = .848$ and $\hat{\beta}^{\text{hom}} = .479$, respectively.[45] In addition, the two subsamples have different final production skill shares. Sectors using more differentiated inputs are on average more skill intensive ($\bar{h}^{\text{diff}} = .286$ vs. $\bar{h}^{\text{hom}} = .245$). This is what we expect, given that differentiated inputs incorporate more product innovation. However, the difference in final production skill shares could also be due to different endowments like high-tech capital or different levels of outsourcing in the two subsamples. To analyze whether this concern is justified, I use the Blinder-Oaxaca decomposition, splitting the mean

[44] To ease graphical exposition, the regressions in Figure 3.2 use equal weights for each sector. The estimated coefficient is very similar when weighted by employment shares, $\hat{\beta} = .558$.

[45] A more detailed analysis, using quintiles of input differentiation κ_i, confirms this result: $\hat{\beta}$ increases with each quintile of κ_i and is highly significant for all except the first one.

outcome differential (predicted $\bar{h}^{\text{diff}} - \bar{h}^{\text{hom}}$) into one part that is due to differences in endowments in the two subsamples, one part that is due to differences in coefficients (after accounting for fixed effects), and a third part that is due to interaction between coefficients and endowments. This decomposition shows that the different final production skill shares in the two subsamples are entirely due to differences in coefficients, while endowments and interaction make small negative (and insignificant) contributions.

Next, I include interaction terms of explanatory variables with input differentiation κ_i.[46] Table 3.7 reports the results, using the three alternative measures for input skill intensity, σ_i (baseline), $\sigma_i^{2\text{d}}$ (excluding inputs from the same 2-digit sectors), and σ_i^w (calculated based on the high-skill wage bill share). The interactions 'input skill intensity' × 'input differentiation' are positive and highly significant, implying that the ITSC grows with the degree of input differentiation. Moreover, the coefficient on input skill intensity (β_1) becomes small and insignificant when the usual controls are included. This indicates that the ITSC is not present for a (hypothetical) sector using only homogenous inputs ($\kappa_i = 0$). To see this, note that the marginal effect of input skill intensity on final production skills is given by $\partial h_i / \partial \sigma_i = \beta_1 + \beta_2 \kappa_i$. On average, this effect is slightly larger than above, where input differentiation was not controlled for.[47]

Table 3.7: Interaction of input skill intensity with input differentiation. Dependent variable is h_{it}.

Input skill measure	σ_i			$\sigma_i^{2\text{d}}$	σ_i^w
	(1)	(2)	(3)	(4)	(5)
Input skill intensity (β_1):	.293*	.118	.046	-.071	.018
σ_i / $\sigma_i^{2\text{d}}$ / σ_i^w	(.152)	(.154)	(.158)	(.240)	(.219)
Inp. skill intensity × inp. differentiation (β_2):	1.118***	1.147***	1.284***	1.325***	1.213***
$\sigma_i \times \kappa_i$ / $\sigma_i^{2\text{d}} \times \kappa_i$ / $\sigma_i^w \times \kappa_i$	(.317)	(.334)	(.312)	(.449)	(.389)
Implied coefficient: $\hat{\beta} = \hat{\beta}_1 + \hat{\beta}_2 \bar{\kappa}$	.907***	.747***	.751***	.657***	.684***
Controls	no	yes	yes	yes	yes
Sector fixed effects	yes	yes	yes	yes	yes
Time fixed effects	yes	yes	no	yes	yes
R^2	.97	.98	.98	.98	.98
R^2 (within)	.53	.59	.58	.58	.59
Observations	2148	2089	2089	2083	2089

Notes: Clustered standard errors (by sector) in parentheses. Key: *** significant at 1%; ** 5%; * 10%. All regressions and the mean $\bar{\kappa}$ are weighted by sectors' average share in total manufacturing employment 1967-92. Controls include: Structures per worker (k^{struct}), equipment per worker (k^{equip}), high-tech capital (HT/K), R&D intensity ($R\&D_{lag}$), and outsourcing (OS^{narr}), as well as their interactions with input differentiation: $k^{\text{struct}} \times \kappa_i$, $k^{\text{equip}} \times \kappa_i$, $HT/K \times \kappa_i$, $R\&D_{lag} \times \kappa_i$, and $OS^{\text{narr}} \times \kappa_i$. Weighted average input differentiation is $\bar{\kappa} = .549$.

[46] Because the framework analyzed here involves complementarity among several explanatory variables, I also interact the control variables with input differentiation. This addresses the concern that the $\sigma_i \times \kappa_i$ interaction alone might capture other effects related to product differentiation. This is the case, for example, if the processing of differentiated intermediates is more R&D intensive, or if outsourcing is more pronounced for differentiated inputs, influencing skill demand through these channels. Input differentiation κ_i is not included in the regressions, as it is captured by sectoral fixed effects.

[47] An interesting and robust finding is that the interaction term 'high-tech capital' × 'input differentiation' is negative and highly significant, while the coefficient on 'high-tech capital' is significantly positive and of the same magnitude (not reported in Table 3.7). Therefore, high-tech capital explains much of the skill demand in sectors using homogenous inputs, but little in sectors using differentiated inputs.

Productivity and ITSC

Now I turn to the relationship between productivity and skills. Because of the well-documented innovation-skill complementarity, we expect sectors with a high proportion of skilled workers to be more productive. However, this holds only for the right mix of complementary inputs [Milgrom and Roberts 1990, Bresnahan et al. 2002]. When skilled workers meet an environment without the potential for production improvements, their innovative potential is wasted. On the other hand, computers or innovative intermediates are squandered when there are no skills to handle them.[48] It is only when skills meet an innovative environment that ideas and productivity flourish. Following this argument, I examine the interaction between input-embedded skills σ_{it}, reflecting innovative intermediates, and final production skills h_{it} in regressions with productivity measures as dependent variable. I run the following regression, expecting a positive coefficient on the interaction term.

$$prd_{it} = \alpha_i + \alpha_t + \beta_1 h_{it} + \beta_2 \sigma_{it} + \beta_3 h_{it} \times \sigma_{it} + \gamma \mathbb{Z}_{it} + \varepsilon_{it} \tag{3.4}$$

where prd_{it} denotes productivity, measured by value added per worker (in natural logarithm) or alternatively by total factor productivity (TFP).[49] $\mathbb{Z}_{it}$ stands for the controls used above, and also includes the interactions of high-tech capital with h_{it} and σ_{it}. As always, sector and time dummies (α_i, α_t) are included, and ε_{it} denotes the error term. The results are presented in Table 3.8.

Table 3.8: Productivity and skills. Dependent variable is productivity.

Productivity measure:	ln(value added per worker)					TFP	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Skilled labor share: h_i	.006		.058	.834***	-1.904***	-2.421***	-2.577***
	(.286)		(.293)	(.179)	(.693)	(.910)	(.899)
Input skill intensity: σ_i	.672	.675		-.025	-1.840**	-2.455**	-1.982*
	(.503)	(.535)		(.382)	(.812)	(.981)	(1.061)
Interaction: $\sigma_i \times h_i$					6.748***	8.824**	9.600***
					(2.089)	(3.610)	(3.567)
Controls	yes	yes	yes	yes	yes	yes	yes
Interaction Controls	no	no	no	no	yes	yes	yes
Sector fixed effects	yes	yes	yes	no	yes	yes	yes
Time fixed effects	yes	yes	yes	yes	yes	yes	no
R^2 (within)	.97	.97	.97	.90	.97	.16	.11
Observations	2089	2089	2089	2089	2089	2089	2089

Notes: Clustered standard errors (by sector) in parentheses. Key: *** significant at 1%; ** 5%; * 10%. All regressions are weighted by sectors' average share in total manufacturing employment 1967-92. Controls include: Structures per worker ($\ln(k^{\text{struct}})$), Equipment per worker ($\ln(k^{\text{equip}})$), high-tech capital (HT/K), R&D intensity ($R\&D_{lag}$), and outsourcing (OS^{narr}). Interaction controls include: $h_i \times HT/K$ and $\sigma_i \times HT/K$.

As column 1 shows, neither input-embedded nor final production skills correlate significantly with productivity. Columns 2 and 3 verify that this is not a consequence of collinearity between σ_i and h_i;

[48] Massive inflows of modern Western capital to the Polish economy in the early 1970's failed to raise industrial productivity – partially due to the lack of technical personnel [Terrell 1992].

[49] I use the 5-factor TFP index (1987=1) from the NBER Manufacturing Industry Database. See Bartelsman and Gray (1996) for a documentation. Results are also very similar when using the natural logarithm of shipments per worker to measure productivity.

neither is significant by itself. Column 4 drops industry fixed effects in order to exploit cross-sectoral variation. In this case, h_i is highly significant, but subject to the concern that unobserved sector-specific characteristics drive both productivity and skill demand. Next, the significantly positive interaction term $\sigma_i \times h_i$ in column 5 explains why σ_i and h_i alone are insignificant: innovative inputs, reflected by their embedded skills, raise productivity only if they are combined with skilled labor to process them. The marginal effect of input skill intensity on productivity is given by $\partial prd_{it}/\partial\sigma_{it} = \beta_2 + \beta_3 h_{it}$, with the weighted average of h_{it} equal to .279, and the corresponding 10$^{\text{th}}$ and 90$^{\text{th}}$ percentile given by .147 and .420, respectively.[50] Therefore, a 1% increase in σ_{it} *lowers* value added per worker by 0.9% when industry i has few skilled workers (10$^{\text{th}}$ percentile), leaves it unchanged in an average industry i, and *raises* value added per worker by 0.9% if i employs many skilled workers (90$^{\text{th}}$ percentile). This finding provides further support for product innovation as the ITSC channel: Skills in intermediate and final production together foster innovation and raise productivity. Finally, columns 6 and 7 confirm that this result is neither an artifact of the chosen productivity measure (it is obtained when using TFP, as well) nor dependent on the inclusion of time dummies.

3.4.4 The ITSC's Contribution to Skill Upgrading

So far, we have seen that the correlation between input skill intensity and the skilled labor share in final production is highly significant and robust to the inclusion of various controls. We have interpreted this finding as evidence for an intersectoral technology-skill complementarity. Next, I turn to the importance of the ITSC for skill demand increases. I present a framework that has been used to estimate the impact of trade and technological change on the demand for skilled labor. The underlying idea is that structural variables like R&D intensity, computer capital, or input skill intensity can shift the production function and therefore the optimal choice of skilled versus unskilled labor. Because some structural variables are arguably endogenous, I use instruments in the corresponding estimation. My results suggest that the ITSC's contribution to skill upgrading in U.S. manufacturing is large – in the same range or even above computers and other high-tech capital.

A labor demand framework

In the following I derive a labor demand regression from a model with sector-specific technologies that are influenced by structural variables.[51] Feenstra and Hanson (1999) argue that labor demand shifting structural variables comprise fixed capital, computing equipment, and outsourcing (reflecting imported input prices). I add R&D intensity and input skill intensity to this list. The former has been identified as an important determinant of skill demand (Machin and van Reenen, 1998), while the inclusion of the latter is motivated by the empirical evidence presented above. Input skill intensity σ_i proxies for innovation or complexity embedded in a sector's intermediates. Because firms need skilled workers to handle innovative intermediates, we expect higher σ_i to go hand in hand with more demand for skilled labor in sector i.

[50] While prd_{it} is specified in logs, σ_{it} and h_{it} are already percentages. Thus, the marginal effect can be interpreted as the elasticity of value added per worker with respect to input skill intensity.

[51] For a more detailed exposition see Katz and Murphy (1992) and Feenstra (2004, ch. 4).

The production function in sector i takes the form $Y_i = F_i(H_i, L_i;\ \sigma_i, \mathbb{Z}_i)$ where structural variables, σ_i and $\mathbb{Z}_i$, are fixed in the short run, while skilled and unskilled labor, H_i and L_i, are chosen optimally. Consequently, a firm in sector i minimizes its wage bill $w_H H_i + w_L L_i$ subject to the corresponding production technology, taking as given high-skill and low-skill wages, w_H and w_L, as well as input skill intensity σ_i and other structural variables $\mathbb{Z}_i$. This yields the short-run cost function: $C_i(w_H, w_L; \sigma_i, \mathbb{Z}_i, Y_i)$. Next, we need to choose a functional form for the cost function. The translog cost function is a convenient choice, as it imposes no a-priori restrictions on elasticities of substitution and returns to scale.[52] We then use Shephard's Lemma, which states that the derivative of the cost function with respect to w_H gives the demand for skilled labor, H_i. This final step is the centerpiece of the demand framework – it enables us to analyze factor demand by examining the properties of the first derivative of the cost function. As shown in Feenstra (2004, ch. 4), we obtain the estimation equation

$$\triangle h^w_{it} = \alpha + \beta \triangle \ln \sigma_{it} + \phi \triangle \ln \mathbb{Z}_{it} + \gamma \triangle \ln Y_{it} + \delta \triangle \ln \left(\frac{w_{H,t}}{w_{L,t}} \right) \tag{3.5}$$

where $h^w_i = (w_H H_i)/(w_H H_i + w_L L_i)$ is the wage bill share of skilled (white-collar) labor. This equation says that the relative demand for skilled labor, represented by its cost share, depends on the structural variables σ_i and $\mathbb{Z}_i$, and on the relative wage. Intuitively, for a given relative wage, structural variables shift the relative demand for the two types of labor, as captured by the coefficients β and ϕ.

Contributions without input skill intensity

The first specification of Table (3.9) presents an estimation of (3.5), following the strategy outlined in Feenstra and Hanson (1999). Input-skill intensity is not yet included in the regression. The structural variables $\mathbb{Z}_i$ comprise all previously used drivers of skill demand (see Table (3.4)). In addition, (3.5) implies that we also have to control for the real value of shipments, $\ln Y_{it}$, and the relative wage $(w_{H,t})/(w_{L,t})$. The latter captures cross-industry variation in wages, for example due to quality variation of workers.[53] Multiplying each regression coefficient by the 1967-92 change in the corresponding variable (shown in the first column) gives each structural variable's contribution to the increase in the white-collar wage share. If we divide this number by the total change in the white-collar wage bill share 1967-92 (0.0727) we obtain percentage contributions. In the first specification, high-tech capital contributes about 12% to overall skill upgrading, and outsourcing (broad and narrow) delivers another 17%. Both numbers are in the ranges documented by Feenstra and Hanson (1999) for the period 1979-90. While the coefficient of R&D intensity is large, its contribution to the increasing white-collar wage share is not. This is because R&D intensity itself increased relatively little. Overall capital is roughly skill neutral, with the positive contribution of structures offsetting the negative impact of equipment.

Endogeneity of input skill intensity

Before including σ_i in regression (3.5), we have to carefully discuss its endogeneity. In the intersectoral complementarity framework presented in the previous sections, causality could run in either direction –

[52] See Kim (1992) for a general treatment of the translog function. In order to derive the estimation equation below, we have to assume that the translog cost function is homogenous of degree one in wages.

[53] Results are very similar when relative wages are dropped from the regressions.

Table 3.9: Contribution to skill upgrading. Dependent variable is the white-collar wage bill share h^w_{it}.

		(1)		(2)		(3)	
	Change '67-'92	Regres-sion	Contri-bution	Regres-sion	Contri-bution	Regres-sion	Contri-bution
Input skill intensity: $\triangle\sigma_i$ *instrumented*	.0546			.485*** (.098)	36.4%	.499*** (.093)	37.5%
Structures per worker: $\triangle k^{\text{struct}}$	.0120	.281 (.220)	4.7%	.196 (.140)	3.2%	.190 (.129)	3.1%
Equipment per worker: $\triangle k^{\text{equip}}$	.0383	-.0679 (.110)	-3.6%	-.0858 (.082)	-4.5%	-.0741 (.070)	-3.9%
High-Tech capital: $\triangle HT/K$ *instrumented in (3)*	.0469	.189* (.115)	12.2%	.299*** (.097)	19.3%	.423** (.190)	27.3%
R&D intensity $\triangle R\&D$ *instrumented in (3)*	.0108	.442*** (.164)	6.6%	.221 (.134)	3.3%	.26 (.242)	3.9%
Outsourcing: $\triangle OS^{\text{narr}}$	.0470	.103** (.043)	6.7%	.0838** (.037)	5.4%	.0816** (.038)	5.3%
Outs.: $\triangle(OS^{\text{broad}} - OS^{\text{narr}})$	.0571	.136*** (.047)	10.7%	.0508 (.056)	4.0%	.0358 (.049)	2.8%
Total Contribution:			37.2%		67.1%		75.9%
Additional Controls:							
Real shipments: $\triangle \ln(Y)$		-.012* (.007)		-.012*** (.004)		-.010** (.005)	
Relative wage: $\triangle \ln(w_H/w_L)$		.101*** (.015)		.118*** (.012)		.115*** (.012)	
Observations		1731		1402		1402	
First stage regressions: ‡							
F-test for significance of IV for σ_i				43.6		35.4	
Instrumented control variables:						HT/K, $R\&D$	
individual F-test for IV						51.5, 15.3	
p-value overidentifying restrictions				.78		.79	
Stock and Yogo weak IV F-statistic				43.5		15.2	
Critical value for highest quality IV				19.9		17.8	

Notes: The first column gives the change of each variable's weighted (by industry wage bill) average over the period 1967-92. The change in the dependent variable h^w is .0727. All regressions are run in 5-year changes and are weighted by sectors' average share in the manufacturing wage bill 1967-92. Standard errors (in parentheses) are robust to arbitrary heteroskedasticity and intra-sector correlation. Key: *** significant at 1%; ** 5%; * 10%. Regressions (2) and (3) are estimated using two-step feasible efficient GMM. 'Contribution' gives the proportion of the observed change in h^w explained by the respective variable.

‡ Instruments are the 5- and 10-year lags of each instrumented variable. In addition, $\triangle\sigma_i$ is instrumented with the 5-year lag of $\triangle\mathbb{Z}_{j\neq i}$ (see text). $\triangle R\&D$ also uses its 15-year lag.

from upstream to downstream skill intensity (σ_i to h_i), or the other way around.[54] But now we treat σ_i as a structural variable that can shift the demand for skilled labor in downstream sectors. We must therefore find instruments that explain innovation and skill upgrading at the upstream level but do not have an impact on downstream skill demand other than through intermediate linkages. To derive candidates for such instruments, I restate the ITSC in a simultaneous equations model:

$$h_{it} = \alpha_{1,i} + \beta_1 \sigma_{it} + \gamma_1 \mathbb{Z}_{it} + \varepsilon_{1,it} \tag{3.6}$$

$$\sigma_{it} = \alpha_{2,i} + \beta_2 h_{it} + \gamma_2 \mathbb{Z}_{j \neq i,t} + \varepsilon_{2,it} \tag{3.7}$$

The first equation represents the upstream-downstream direction of the ITSC, estimating the impact of input skill intensity on final production skill demand, where $\mathbb{Z}_{it}$ are the usual control variables. The second equation describes the opposite causal direction – from final producers i to intermediate suppliers $j \neq i$.[55] Control variables that affect the skill demand in intermediate production, $Z_{j \neq i,t}$, are constructed in a similar fashion as σ_{it}:

$$Z_{j \neq i,t} = \sum_{j \neq i} \bar{a}_{ij} Z_{jt} \tag{3.8}$$

For example, let Z_{jt} be outsourcing of intermediate supplier j. We expect this variable to affect $\sigma_{it} = \sum_{j \neq i} \bar{a}_{ij} h_{jt}$ through its impact on h_{jt}. The same holds for all suppliers $j \neq i$. In this example $Z_{j \neq i,t}$ thus represents weighted average outsourcing of sector i's suppliers. The same methodology applies to computer and high-tech capital as well as R&D intensity. All are summarized as $\mathbb{Z}_{j \neq i,t}$.[56]

Deriving the reduced form for σ_{it} from (3.6) and (3.7) gives the first stage of an instrumental variable (IV) regression, with $\mathbb{Z}_{j \neq i,t}$ being the instruments for σ_{it}:

$$\sigma_{it} = \tilde{\alpha} + \tilde{\gamma_1} \mathbb{Z}_{it} + \tilde{\gamma_2} \mathbb{Z}_{j \neq i,t} + \tilde{\varepsilon}_{it} \tag{3.9}$$

This equation suggests that we could use $\triangle \mathbb{Z}_{j \neq i,t}$ to instrument for $\triangle \sigma_{it}$ in (3.5). I use the 5-year lags, $\triangle \mathbb{Z}_{j \neq i,t-5}$, in order to alleviate the concern that downstream innovation and skill upgrading might lead to more computer use or R&D in supplying upstream industries.[57] The exclusion restriction is that instruments $\triangle \mathbb{Z}_{j \neq i,t-5}$ influence $\triangle \sigma_{it}$ but are uncorrelated with $\triangle h^w_{it}$ once we control for $\mathbb{Z}_{it}$ in the second stage, i.e., in (3.5). For this restriction to hold we have to assure that, first, lagged changes in high-tech capital, outsourcing, or R&D at the upstream level influence upstream skill demand and, second, do only have an impact on downstream skill demand through the (intersectoral) technology-skill complementarity. The first part is well-founded, as my own and other previous findings in the literature show. Note that variations in high-tech capital or R&D across sectors may also capture variations in

[54] As a first pass at the issue, I use a Granger causality test. The usual cavaets apply – time precedence and causality are two distinct concepts. I find Granger causality in both directions: stronger in the upstream-downstream direction, where the coefficient on lagged σ_i is .179 (.054); and weaker in the opposite direction with a coefficient on lagged h_i of .033 (.015). All lags are 5 years; both regressions include all lagged control variables used in Table 3.4, sectoral dummies, and the lag of the left-hand side variable.

[55] These two equations can be used to quantify the bias that arises when we interpret the OLS coefficient β from equation (3.3) as a causal influence of σ_{it} on h_{it}, not taking into account the reverse relationship. The covariance between σ_{it} and $\varepsilon_{1,it}$ is given by $\beta_2/(1-\beta_1\beta_2)Var(\varepsilon_{1,it})$, and the corresponding bias in β is equal to this covariance divided by $Var(\sigma_{it})$. The Granger causality test suggests that the feedback from h_{it} to σ_{it} is small in comparison to the opposite direction. We thus expect that β_2 is small, which yields a small positive bias of the OLS coefficient in (3.3).

[56] The set $\mathbb{Z}_{j \neq i}$ that I use in the IV regressions comprises $(HT/K)_{j \neq i}$, $(OCAM/K)_{j \neq i}$, $OS^{\text{narr}}_{j \neq i}$, $OS^{\text{broad}}_{j \neq i} - OS^{\text{narr}}_{j \neq i}$, and $(R\&D)_{j \neq i}$.

[57] Recall that the manufacturing industry panel has 5-year intervals. Using the contemporaneous $\triangle \mathbb{Z}_{j \neq i,t}$ gives similar results.

(unobserved) innovative activity. This poses no problem for the instruments – to the contrary: it is in line with the innovation - skill channel, which is at the core of the ITSC. Although less evident, it is reasonable to argue that the instruments also fulfill the second requirement. More computers and R&D in an upstream sector can drive innovation and skill upgrading there, leading to downstream sectors demanding more skills in order to process the innovative intermediates. But upstream computers requiring downstream skills for reasons different from innovation-skill complementarities is harder to maintain – especially because any intersectoral computer-compatibility channel would be captured by computers showing up as a control variable in the second stage regression (3.5).

A final concern regarding the instruments is that high-tech capital, outsourcing, or R&D can be correlated across upstream and downstream industries, especially if the production chain involves similar industries. This could lead to $\triangle\mathbb{Z}_{j\neq i,t-5}$ influencing $\triangle h_{it}^{w}$ because both correlate with $\triangle\mathbb{Z}_{it}$. Including the downstream variables $\triangle\mathbb{Z}_{it}$ in the second stage regression (3.5) controls for this channel. Although we are able to alleviate the most important concerns, it is important to mention that the instruments are not completely satisfactory. Endogeneity remains a concern if three things come together: unobserved shocks or innovations hit similar industries, these industries are linked through intermediates, and the shocks influence skill demand and the $\mathbb{Z}$-variables over a long horizon (>5 years). Empirically, we have the means to shed light on this concern using overidentification restrictions. The results are encouraging (see below).[58]

Endogeneity of control variables

Having addressed the endogeneity of σ_{it}, we now turn to the same concern for the other structural variables $\mathbb{Z}_{it}$. Most importantly, endogeneity is an issue for high-tech capital as well as for R&D intensity. To tackle the potential bias, I use an approach outlined in Wooldridge (2002, ch. 11). Under sequential exogeneity, we can use lagged levels of $\mathbb{Z}_{it}$ as instruments for $\triangle\mathbb{Z}_{it}$, which gives consistent estimates and is similar in spirit to Arellano and Bond (1991). Sequential exogeneity implies that if we run regression (3.5) in levels, then after the structural variables ($\mathbb{Z}_{it}$ and σ_{it}) and sectoral fixed effects have been controlled for, no past values of $\mathbb{Z}_{it}$ or σ_{it} affect the expected value of h_{it}^{w}. To see whether this holds, I include the 5- and 10-year lags of all structural variables in (3.5), estimated in levels. None is significant at the 10% level, with the exception of the 10-year lag of $R\&D$. This suggests that sequential exogeneity is a reasonable assumption for the structural variables in my demand framework. As suggested in Wooldridge (2002, ch. 11), I thus use the 5- and 10-year lags of high-tech capital and R&D intensity to instrument for the contemporaneous 5-year changes. Sequential exogeneity also delivers the 5- and 10-year lags of σ_{it} as additional instruments for $\triangle\sigma_{it}$.[59]

[58] An empirical analysis that takes a further pass at the endogeneity issue is in the making. This project combines changes in Argentinian tariffs at the detailed 4-digit level with firm-level workforce characteristics at the same level of detail. Following sector-specific drops in tariffs, Argentinian firms increase their imports of US intermediates. In sectors where this leads to an increase of input skill intensity, we expect skill demand in Argentinian final production to rise.

[59] My estimation results do not depend on whether or not I use these additional instruments, but they contribute to instrument quality and provide additional overidentification restrictions (see below).

Results with input skill intensity

Following this extended discussion, it is time to turn to the estimation results. The second specification in Table 3.9 adds instrumented input skill intensity to the regression. The corresponding coefficient is highly significant and smaller than in the OLS specification.[60] This makes sense, given that we expect an upward bias of OLS estimates (see footnote 55). Input skill intensity contributes over one third to the overall increase in the white collar wage bill share in US manufacturing – about as much as the upper bound of previous estimates for the contribution of computers. Interestingly, the other structural variables remain largely unchanged when adding $\triangle\sigma_{it}$ to the regression, which suggests that input skill intensity is not merely picking up explanatory power from other variables. As the number of observations reflects, the choice of instruments – using lagged changes $\triangle\mathbb{Z}_{j\neq i,t-5}$ as instruments for $\triangle\sigma_{it}$ – loses an additional time period (with one already lost due to first differencing). The instruments for $\triangle\sigma_{it}$ are highly significant – the corresponding F-statistic of the exclusion hypothesis is well above the rule of thumb threshold of 10 recommended by Staiger and Stock (1997) to avoid weak instrument concerns. The additional test of weak instruments based on Stock and Yogo (2002) confirms this result. This test becomes especially useful in models with more than one endogenous variables and is discussed in more detail below. Since the number of instruments is larger than one, we can test for their endogeneity using the Sargan-Hansen test of overidentifying restrictions. The corresponding p-value is .78. We therefore do not reject instrument exogeneity.

Next, I turn to specification 3, instrumenting for several endogenous structural variables. The Staiger and Stock (1997) rule of thumb for avoiding weak instruments refers to models with one endogenous variable. In models with two or more endogenous variables, instruments can be weak despite being very significant in each first-stage regression. This is because endogenous explanatory variables predicted by the instruments may be close to collinear, which makes it difficult to separate the effect of each individual one. Stock and Yogo (2002) provide a framework that allows testing the hypothesis of weak instruments in this case. The null hypothesis is that instrument quality is below one of four levels. The last row of Table 3.9 reports the critical value for the highest quality level, corresponding to a maximum IV bias of 5% because of weak instruments. The Stock and Yogo framework allows for models with up to three endogenous variables. Therefore, I first instrument (in addition to σ_{it}) for those two controls for which endogeneity is the most serious concern: high-tech capital and R&D intensity.[61] The results for all structural variables, including input skill intensity, are very similar to the previous specification – with the exception of high-tech capital that now has a larger coefficient. The p-values for the overidentification test is again well above the rejection level. Finally, instruments are close to the highest quality level according to the Stock and Yogo test.[62] Altogether, the results reported in Table 3.9 suggest that the ITSC is very important for explaining skill upgrading in US manufacturing. Its contribution appears to be in the same order of magnitude, or even larger, than high-tech capital.

[60] See, in particular, regression (3) in Table 3.6, which also uses h^w_{it} as dependent variable.

[61] R&D intensity can be instrumented with one more time lag without using an additional time period of observations, because the R&D data include 1963.

[62] The critical value for the second quality level, corresponding to a maximum IV bias of 10%, is 10.01.

3.5 A Sketch Model

This section integrates my empirical findings into the analytical SBTC framework. The standard setup has two types of labor in a CES production function, producing one final good.[63] I add intermediate input linkages and skill-complementarity across many sectors, as motivated by the empirical evidence presented above. In order to concentrate on the main mechanism, I present a static model, abstracting from intertemporal dynamics and endogenous skill supply. The economy is composed of $i = 1, ..., N$ sectors, each producing a specific good, or variety i. The number of sectors is fixed. Within each sector, a multiplicity of firms operates under perfect competition and constant returns. I focus on a representative firm for each sector i, making zero profits. Each good i is used for final consumption and as intermediate input in sectors $j \neq i$ with constant input shares. This Leontief technology is at the heart of input-output tables, and section 3.4.2 has shown that constant input shares are a reasonable assumption.

The economy is populated by L low skilled individuals, working in production, and H high-skilled individuals that coordinate production and handle innovative intermediate inputs. The skill intensity of inputs is defined as the weighted average share of high-skilled workers employed in their production, resembling the empirical part of the paper. High-skilled workers are relatively more productive in processing skill intensive inputs. This setup is similar in spirit to Kremer's (1993) O-Ring theory. Kremer assumes that production involves the completion of n tasks, each performed by a worker of skill level q_i in one and the same firm. Output is proportional to $\prod_{i=1}^{n} q_i$, implying a strong complementarity of workers' skill levels. In this framework, a high skilled worker performing task i is most productive in firms that employ high-q workers in all other tasks, too. The model presented here can be thought of as a multi-sector version of the O-Ring theory. Kremer's tasks are my intermediate inputs – final products contain intermediates from various sectors instead of being entirely manufactured in one firm.[64] Innovations and quality of skilled workers are embedded in the goods they produce. In my setup, Kremer's *within*-firm skill complementarity works its wonders *across* firms along the production chain. High skilled workers in sector N are the more productive relative to the unskilled, the more innovative their inputs are, i.e., the more skills are embedded in the $N-1$ input varieties that they process.

Another related model endogenizes the direction of technical change [Acemoglu 1998, 2002 and 2007, Acemoglu and Zilibotti 2001]. Therein, an increasing number of skilled workers implies a larger market and demand for skill-complementary technologies, inducing skill-biased technological change. However, this channel lacks empirical support, as it is hard to pin down a robust relationship between demand factors and R&D intensity [Cohen and Levin 1989].[65] In a more recent contribution, Ngai and Samaniego (2007) find that neither TFP growth nor R&D intensity are related to demand factors in equilibrium, arguing that technical progress is largely a supply-driven phenomenon. They draw this conclusion from a calibrated multi-sector model of productivity growth with knowledge generation and spillovers as the driving factors. This is similar in spirit to my model, where innovations and skill demand are also supply driven. However, neither the model of Ngai and Samaniego nor the one pioneered by Acemoglu feature intersectoral linkages or skill complementarities across sectors. Furthermore, in Acemoglu's setup a relative increase in the amount of skill-complementary technologies yields decreas-

[63] See Card and DiNardo (2002) for a review of the standard SBTC framework.

[64] The infamous O-Ring itself is not a product of the space vehicle industry (SIC 3761), but of the gaskets, packing, and sealing devices sector (SIC 3053).

[65] Some studies argue explicitly against demand-driven innovation [Nelson and Winter 1977].

ing incentives to develop more of them, since their relative price falls.[66] In my model, increasing skill intensity in one sector *augments* the skill bias in other sectors connected through intermediate linkages.

3.5.1 Production and Consumption

There are N types of goods produced in this economy, each by a representative firm in its corresponding sector i. A producer of good i employs low-skilled labor L_i, high-skilled labor H_i, and an aggregate of intermediate inputs $\mathcal{X}_i$, specified in more detail below. Output of good i is given by

$$Y_i = A_i \left[\gamma_i \left[e^{\,\phi_i \sigma_i} \, H_i \right]^{\frac{\epsilon-1}{\epsilon}} + (1-\gamma_i) \left[L_i \right]^{\frac{\epsilon-1}{\epsilon}} \right]^{\frac{\epsilon}{\epsilon-1}\alpha} (\mathcal{X}_i)^{1-\alpha} \tag{3.10}$$

where $\epsilon > 0$ is the elasticity of substitution between the labor inputs, γ_i is a sector-specific technology parameter, and α is the share of value added (or aggregate labor) in production. Finally, σ_i denotes the skill intensity of intermediate inputs that enter the production of i, and ϕ_i reflects the strength of the ITSC. I do not specifically model innovation, but rather assume that skilled labor H_i performs innovation and is needed to process innovative intermediates. This is a shortcut, aimed at providing a simple calibratable model. A micro-founded model is in the works [Voigtländer 2008]. If $\phi_i > 0$, *overall* productivity of sector i increases in input skill intensity σ_i, which reflects the spillover of innovations embedded in intermediates produced by skilled workers. Moreover, when $\phi_i > 0$, the *relative* productivity of high-skilled workers increases with σ_i, reflecting skill complementarities along the production chain. Therefore, a sector purchasing skill intensive intermediates will employ relatively more skilled workers. If $\phi_i = 0$ we are back to a standard SBTC production function, in a setup with intermediate inputs.

Each sector i uses the products from all sectors $j \neq i$ as intermediate inputs. To keep matters simple, I assume that intermediates enter final production (3.10) according to a Leontief technology:

$$\mathcal{X}_i = \min_{j \neq i} \left\{ \frac{1}{a_{ij}} X_{ij} \right\} \tag{3.11}$$

where X_{ij} is the amount of input j used in the production of good i, and $a_{ij} \in (0,1)$ is the corresponding input requirement. High a_{ij} indicate that much of input j is needed in the production of product i. Sectors do not use their own output as intermediate: $a_{ii} = 0$, but use a positive amount of all others: $a_{ij} > 0$, $\forall j \neq i$; and a_{ij} is normalized such that $\sum_{j \neq i} a_{ij} = 1$. Let the fraction $x_{ij} \equiv X_{ij}/a_{ij}$ denote the effective units of input j. When optimizing production, the representative firm i chooses the same amount of each effective input j, such that $x_{ij} = \overline{x}_i, \ \forall j$. Consequently, the total amount of input i used by sector j is given by

$$X_{ij} = a_{ij} \overline{x}_i \tag{3.12}$$

where the effective amount of each input in sector i, $\overline{x}_i$, is determined in the optimization of production (3.10), with $\mathcal{X}_i = \overline{x}_i$. A convenient feature is that $\overline{x}_i$ also gives the total amount of intermediates used, $\sum_{j \neq i} X_{ij} = \overline{x}_i$. Equation (3.12) implies that the share of input j in sector i is given by $X_{ij} / \sum_{j \neq i} X_{ij} = a_{ij}$. The final piece of the model's production side is the skill intensity of inputs, which is defined in

[66] The overall strength of the skill bias results from a trade-off between this price effect and the market size effect, with the elasticity of substitution between skilled and unskilled labor playing a crucial role. This parameter is of secondary importance for my results.

concordance with the empirical analysis:

$$\sigma_i = \sum_{j \neq i} a_{ij} h_j \tag{3.13}$$

where h_j is the skilled labor share employed in the production of input j. Thus, $\sigma_i \in [0, 1]$ represents the weighted average share of skilled workers employed in the production of all intermediate inputs used in sector i.

All agents have the same preference structure, independent of their skill level. Skill-specific wages w_L and w_H are the only source of income. There is no investment. A representative consumer draws utility from the consumption c_i of all N goods according to the Cobb-Douglas preferences

$$u\left(\{c_i\}_{i=1}^N\right) = \exp\left(\sum_{i=1}^N \ln c_i\right). \tag{3.14}$$

This formulation of utility is convenient because it delivers constant and equal final expenditure shares in equilibrium.

3.5.2 Linkages, Complementarities, and Multipliers

The economic environment in my model is similar to Jones' (2007) setup involving intermediate linkages and complementarity. There, too, final goods are used for both consumption and as intermediate inputs. Jones needs the assumption that goods are complements in both production and consumption in order to obtain a closed-form solution. My approach achieves this result with a more natural formulation of preferences but the stronger assumption of no input substitutability. Both Jones' and my model deliver a multiplier that reinforces productivity differences and skill demand, respectively. The multiplier channel, however, is different. In Jones' paper, higher intermediate productivity leads to more output, which feeds back into the production of intermediates. The share of intermediate goods in total revenue is therefore crucial for the size of the multiplier. In my approach, the intermediate input share in total output, $1 - \alpha$, is not important for the ITSC. What counts is the average proportion of skills embedded in inputs, σ_i, together with the strength of cross-sectoral complementarity given by ϕ_i. Linkages are only important for granting that sectors process each others' output. They are necessary, but not sufficient for skill complementarities across sectors. If $\phi_i \sigma_i = 0$, there is no intersectoral skill complementarity despite the existence of intermediate linkages. Provided that $\phi_i > 0$, my model delivers a skill demand multiplier. Suppose that H_j increases relative to L_j, for example because of an innovation in sector j that requires more skilled labor. Skill upgrading in sector j increases σ_i for all $i \neq j$, which leads to higher productivity of skilled workers and thus augmented skill demand in these sectors, as well. The consequence is a virtuous circle of skill upgrading in the whole economy. Appendix A.3 derives the multiplier effect formally.

3.5.3 Optimization and the Symmetric Case

Firms take factor and goods prices as given and choose L_i, H_i, and $\overline{x}_i$ to maximize profits from production (3.10) subject to (3.11) - (3.13). $\mathcal{X}_i$ in (3.10) is replaced by $\overline{x}_i$ because of the Leontief technology related to intermediate inputs. The total cost of intermediates is $\sum_{j \neq i} p_j X_{ij}$, with X_{ij} given by (3.12). A representative firm in sector i optimizes

$$\max_{\{L_i, H_i, \overline{x}_i\}} p_i Y_i - w_L L_i - w_H H_i - \sum_{j \neq i} p_j a_{ij} \overline{x}_i \tag{3.15}$$

where p_i is the price of good i. A convenient implication of the Leontief technology is that firms do not adjust intermediate input proportions if input skill intensities change, that is, firms take σ_i as given. Setting the ratio of the two labor types' marginal product equal to the ratio of their wages and rearranging yields the relative demand for skilled labor:

$$\frac{H_i}{L_i} = \left(\frac{\gamma_i}{1-\gamma_i}\right)^{\epsilon} \left(e^{\phi_i \sigma_i}\right)^{\epsilon - 1} \left(\frac{w_L}{w_H}\right)^{\epsilon} \tag{3.16}$$

The relative labor demand is determined by sector-specific characteristics γ_i (including, for example, computer equipment and outsourcing), input skill intensity, and relative wages. This result will become important in the calibration of the model. The remaining steps of calculus for the production side are needed to close the model, but not crucial for the intuition. They are presented in Appendix A.2.

On the demand side, let $c_{L,i}$ and $c_{H,i}$ denote labor-type specific consumption of good i. Low-skilled and high-skilled individuals maximize (3.14) subject to their budget constraints $\sum_{i=1}^{N} p_i c_{L,i} \leq w_L$ and $\sum_{i=1}^{N} p_i c_{H,i} \leq w_H$, respectively. This yields the skill-specific demand functions

$$c_{L,i} = \frac{w_L}{N p_i} \quad \text{and} \quad c_{H,i} = \frac{w_H}{N p_i} \tag{3.17}$$

Let $C_i = L c_{L,i} + H c_{H,i}$ denote total *final* demand, and $X_{\bullet i} = \sum_{j \neq i} X_{ji}$ total *intermediate* demand for good i. We can now specify the three market clearing constraints that the economy faces:

$$L = \sum_{i=1}^{N} L_i \tag{3.18}$$

$$H = \sum_{i=1}^{N} H_i \tag{3.19}$$

and

$$Y_i = C_i + X_{\bullet i}, \; \forall i. \tag{3.20}$$

The first two constraints assume that the economy is endowed with an exogenously given amount of each type of labor, and that both are fully employed. The last market clearing constraint says that each sector's output is completely used up in final consumption and as an intermediate input for other sectors' production.

For expositional reasons, I present only the symmetric case of the model. This is sufficient to explain the main intuition, and more readily compared to the standard SBTC framework. However, heterogeneity of sectors is important in the calibration, as it provides the variation needed to identify the key parameter ϕ.

Definition 1 *The symmetric case of the model is characterized by all sectors having the same technology, that is, $A_i = A$, $\gamma_i = \gamma$, $\phi_i = \phi$, $\forall i = 1, ..., N$; and $a_{ij} = 1/(N-1), \forall j \neq i$.*

The last expression in the definition says that each sector uses the same proportion of all other sectors' products as intermediate inputs. Appendix A.2 shows that in the corresponding symmetric equilibrium the relative wage is given by

$$\frac{w_H}{w_L} = \frac{\gamma}{1-\gamma}\left(e^{\phi h}\right)^{\frac{\epsilon-1}{\epsilon}}\left(\frac{L}{H}\right)^{\frac{1}{\epsilon}} \tag{3.21}$$

where h is the proportion of high-skilled workers in the economy. This result is an extension of the standard expression in the SBTC literature, which is recovered if $\phi = 0$, i.e., in the absence of intersectoral skill complementarities. The empirical evidence presented above argues strongly for $\phi > 0$. In this case, an increase in H relative to L has two effects. First, the standard downward pressure on the relative wage due to the increased relative supply. Second, the ITSC effect, working in the opposite direction: The newly employed skilled workers foster product innovation in their own sectors, which in turn drives innovation in all other sectors and raises the relative productivity of skilled workers. The second effect therefore raises skill demand and the relative wage. Next, I calibrate the model in order to investigate the strength of the ITSC effect and see how the model performs in explaining the observed relative wage trend in U.S. manufacturing.

3.5.4 Calibration

In the symmetric equilibrium shown in (3.21), ϕ represents the average strength of the ITSC in the model economy. In order to calibrate this parameter, I use my panel of manufacturing sectors. First, I derive the relative demand for skilled workers from equation (3.16) in logarithmic form:

$$\ln\left(\frac{H_i}{L_i}\right) = \ln\left(\frac{\gamma_i}{1-\gamma_i}\right) + (\epsilon - 1)\phi_i\sigma_i + \epsilon\ln\left(\frac{w_L}{w_H}\right) \tag{3.22}$$

On the right-hand side, γ_i reflects sector-specific characteristics driving skill demand, i.e., the previously used sectoral fixed effects and control variables. I follow two approaches to deal with the inverse relative wage in (3.22). When estimating (3.22) in levels, I use time-dummies to account for changes in economy-wide relative wages. When run in changes, I include the relative wage in the regression. This also accounts for sector-specific worker quality.[67] A variety of studies pin down the elasticity of substitution between high- and low-skilled labor, ϵ, in the range 1.5 to 2 [Angrist 1995, Ciccone and Peri 2005]. In (3.22) we can only identify the term $\beta_i \equiv (\epsilon - 1)\phi_i$. However, for a given ϵ, ϕ_i can be recovered. Following the empirical findings reported above, we expect $\beta_i \geq 0$ and increasing in the degree of input differentiation in sector i. The identifying regression is:

$$\ln\left(\frac{H_{it}}{L_{it}}\right) = \alpha_i + \alpha_t + \beta_i\sigma_{it} + \gamma\mathbb{Z}_{it} + \varepsilon_{it} \tag{3.23}$$

where α_i and α_t are sector and time fixed effects, $\mathbb{Z}_{it}$ are control variables, and ε_{it} denotes the error term. Note the similarity to regression (3.5), which we derived from the labor demand framework. The left-hand side variable is now the relative demand, rather than the wage share of skilled labor. There are two ways to estimate the economy-average ITSC parameter ϕ. First, identify it directly by constraining $\beta_i = \beta, \forall i$ and weighting by sectoral employment. The corresponding results are shown in column 1 (using OLS) and columns 3, 5, and 6 (using instruments) of Table 3.10. Instruments and control variables are the same as in section 3.4.4; the previous discussion of control variable endogeneity and potential bias applies here, as well. OLS and IV estimates yield similar estimates of the coefficient β.[68] Second,

[67] See footnote 38 for a discussion. I also address the endogeneity of controls as in section 3.4.4.

[68] Since Stock and Yogo's (2002) weak instrument test is only available for up to three endogenous regressors, I instrument (in addtion to σ_i and the relative wage) for the one for which endogeneity is of greatest concern: high-tech capital.

take into account that β_i varies with input differentiation κ_i and include the corresponding interaction: $\beta_1\sigma_i + \beta_2\sigma_i\kappa_i$ (using the interaction terms $\mathbb{Z}_{j\neq i} \times \kappa_i$, to instrument for $\sigma \times \kappa_i$). In this case, the average effect is $\beta = \beta_1 + \beta_2\bar{\kappa}$, where $\bar{\kappa}$ is average input differentiation, weighted by sectoral employment. Columns 2 and 4 show the corresponding results, with the derived coefficient slightly larger as compared to the first method. As reported in the bottom of the table, instruments pass all the relevant tests.[69] Overall, the estimates of β lie in the range 2.2-3.5. I use the IV estimate of column 3, $\beta \approx 3.0$, as a baseline, and also include the lower and upper bounds in the calibration.

Table 3.10: Calibration of the ITSC parameter ϕ. Dependent variable is $\ln(H_i/L_i)$.

	Levels		Changes			
	OLS	OLS	IV	IV	IV	IV
	(1)	(2)	(3)	(4)	(5)	(6)
Input skill intensity (β_1):	2.670***	1.088	3.052***	1.759*	2.543***	2.187***
σ_i	(.629)	(1.095)	(.514)	(.974)	(.464)	(.445)
Inp skill intens. $\times$ inp diff. (β_2):		4.045*		3.184		
$\sigma_i \times z_i$		(2.065)		(2.018)		
Implied coefficient: $\hat{\beta} = \hat{\beta}_1 + \hat{\beta}_2\bar{z}$		3.309***		3.508***		
Relative wage: $\ln(w_{H,i}/w_{L,i})$			-.452***	-.770***	-.467***	-.515***
			(.051)	(.139)	(.053)	(.160)
Controls	yes	yes	yes	yes	yes	yes
Sector fixed effects	yes	yes	no	no	no	no
Time fixed effects	yes	yes	no	no	no	no
R^2 (after FE)	.56	.57				
Observations	2089	2089	1402	1402	1402	1402
First stage regressions: ‡						
F-test for significance of IV for: $\triangle\sigma_i$			39.82	56.82	33.86	29.8
$\triangle\sigma_i \times z_i$				62.52		
Instrumented control variables:					HT/K	HT/K, $\ln(w_H/w_L)$
p-value overidentifying restrictions			.82	.20	.64	.42
Stock and Yogo weak IV F-statistic			39.8	24.0	52.3	9.8
Critical value for highest quality IV			19.9	19.8	18.3	17.4

Notes: Clustered standard errors (by sector) in parentheses. Key: *** significant at 1%; ** 5%; * 10%. All regressions and the mean $\bar{\kappa}$ are weighted by sectors' average share in total manufacturing employment 1967-92. Controls include the following variables: Structures per worker (k^{struct}), Equipment per worker (k^{equip}), High-Tech capital (HT/K), R&D intensity ($R\&D_{lag}$), and Outsourcing (OS^{narr}). In columns 2 and 4, also the interactions of the control variables with input differentiation κ_i are included. Weighted average input differentiation is $\bar{\kappa} = .549$. All variables in (1) and (2) are in levels, while 5-year differences are used in regressions (3) - (6); the latter are estimated using two-step feasible efficient GMM.
‡ Instruments are the 5- and 10-year lags of each instrumented variable. In addition, $\triangle\sigma_i$ is instrumented with the 5-year lag of $\triangle\mathbb{Z}_{j\neq i}$. See section 3.4.4 for further details on instruments.

Figure 3.3 shows the results of the calibrated model, depicting the skill premium given by (3.21). In the absence of other factors driving skill demand (γ_i constant), the model with $\phi = 0$ predicts a sharp decline in w_H/w_L when the high-skill labor share h grows. In the figure, I refer to this as the standard model, meaning a CES production function with skilled and unskilled labor. The ITSC model uses $\phi =$

[69] In column 6, instruments are not of the highest quality, but very close to the second-best level, corresponding to a maximum IV bias of 10% (critical value: 9.85).

3.0, corresponding to $\beta = 3.0$ and $\epsilon = 2$.[70] In the ITSC baseline case, the decline of the skill premium is much more moderate. The same holds for the upper and lower bounds, $\phi = 2.2$ and $\phi = 3.5$. This result is interesting when related to the endogenous SBTC literature pioneered by Acemoglu. Therein, elasticities $\epsilon > 2$ are needed to obtain an increasing skill premium as a response to increasing skill supply. My results suggest that when intersectoral skill complementarities are added to this setup, more realistic elasticities $\epsilon < 2$ will deliver increasing skill premia because the ITSC flattens the aggregate skill demand curve.

The right panel of Figure 3.3 compares the model predictions with the observed skill premium in U.S. manufacturing. While the weighted average share of skilled workers rose from 24.7 to 30.6 percent between 1967 and 1992, the skill premium returned to its previous value of 1.56 after a small initial decline. Following the common convention, I refer to skill bias as the difference between the standard model's predicted decline and the observed stagnation of the relative wage. The calibrated ITSC model explains about half of the observed skill bias. This confirms the empirical importance of the ITSC that we found in section 3.4.

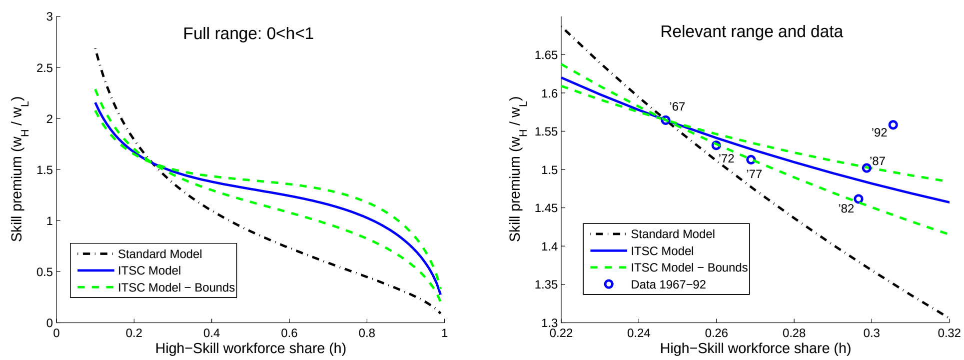


Figure 3.3: Calibrated ITSC model vs. standard model and data

Notes: The data in the right panel represent the weighted average share of white-collar workers in U.S. manufacturing 1967-92, derived from the NBER Manufacturing Industry Database, using total sectoral employment as weights. The parameter γ is normalized such that the model matches the data in 1967. The elasticity of substitution between skilled and unskilled labor is $\epsilon = 2$.

3.6 Conclusions

While intermediate inputs account for more than half of a final product's value, intersectoral linkages have been ignored as a source of skill bias. Existing empirical work on rising wage inequality has failed to account for the full scope of skill upgrading in recent decades. This paper presents strong evidence for an intersectoral technology-skill complementarity (ITSC). The ITSC amplifies initial shocks or innovations that increase skill demand, spreading their impact across sectors. I provide empirical evidence suggesting that the ITSC works through product innovation. The innovative activity of skilled workers in one sector improves products used in many other sectors, stimulating innovation and skill demand along the production chain. The result is a self-enforcing circle of skill upgrading that eventually feeds back

[70] Results are very similar when using $\epsilon = 1.5$ and $\phi = 3.0/0.5 = 6.0$.

into the originating sector. Overall, the ITSC can account for more than one third of the skill upgrading in U.S. manufacturing between 1967 and 1992. The remaining is largely explained by previously suggested within-sector drivers of skill demand, including high-tech capital, R&D intensity, and outsourcing.

To identify this novel mechanism, I construct a measure for the skills embedded in a sector's intermediate inputs. This input skill intensity correlates with final production skills, i.e., skills employed in the further processing of intermediates. The correlation is robust to the inclusion of numerous control variables previously suggested in the SBTC literature, as well as to using a more conservative measure of input skill intensity, discarding linkages between similar sectors. These results are confirmed by an estimation framework that goes beyond the mere correlation, using instruments to account for the bi-directional causality between upstream and downstream skill requirements and the endogeneity of control variables.

The ITSC does not come as a surprise. It combines the well-documented findings of a technology-skill complementarity *within* sectors with technological spillovers *across* sectors. The concept of multipliers due to intermediate linkages is also a well-established one. It has been used in studies explaining productivity differences or rising world trade, but not in the SBTC literature. Two empirical findings suggest that the ITSC works through product innovation performed by skilled workers. First, the ITSC is stronger when involving differentiated intermediates, more readily reshaped by innovative minds. I show that product innovation is more pronounced in sectors producing differentiated goods. Thus, downstream industries using differentiated intermediates purchase relatively more embedded innovation. Constructing a measure of input differentiation, I then provide evidence for a stronger ITSC among sectors linked through differentiated intermediates. Second, productivity regressions show that skills in intermediate and final production complement each other in driving TFP and output per worker. Skill-intensive intermediates go hand in hand with higher productivity only if they meet skills in final production. These findings suggest that upstream skills foster intermediate product innovation, which in turn augments skill demand and productivity in final production.

In order to integrate my empirical findings into the SBTC framework, I extend the standard model featuring skilled and unskilled labor in a CES production function, adding intermediate inputs in a setup with N sectors. Therein, the *relative* productivity of skilled workers can grow with skills embedded in intermediate inputs, reflecting the complementarity of skills along the production chain. Moreover, *overall* productivity increases with input skill intensity, reflecting the innovative activity of skilled workers in intermediate production. An increase in the number of skilled workers has two effects on the skill premium: The standard downward pressure due to increased supply, and an ITSC effect, pushing in the opposite direction. The latter works through the complementarity of skills along the production chain. Once the newly available skills are employed in one sector, they raise the relative productivity of skilled labor in other sectors through intermediate linkages, augmenting skill demand. The calibrated model can account for almost half of the observed skill bias in U.S. manufacturing, confirming the previous estimation results.

The present paper documents the novel stylized fact of an ITSC and applies it to the skill bias of technical change. In addition, the ITSC opens the door for analyzing other important questions from an intermediate-linkage angle. One example is the observed prevalence of North-North trade. Standard models of international trade predict that the skill-abundant North should specialize in skill-intensive production, importing low-skill intensive goods from the South. The ITSC, on the other hand, suggests that skill intensive Northern production requires high-quality skill-intensive intermediates, purchased in the North. Another potential application is the observation that TFP growth rates differ widely and persistently across industries [Ngai and Samaniego 2007]. The ITSC, working through product innovation, can help to explain this fact. Some sectors purchase more innovation embedded in their intermediates than others. Innovative intermediates, in turn, foster final product improvements. Therefore, heterogeneity in intermediate input requirements could lead to persistent variations in sectoral TFP growth. Finally, an important topic for further investigation is whether the ITSC is a broad phenomenon, extending to linkages beyond the manufacturing sector.

3.7 APPENDIX

3.7.1 Appendix A.1 Data Sources and Construction of Variables

Product innovation

Scherer (1982) provides data on R&D expenditures broken down into product and process innovation for 36 manufacturing sectors, broadly equivalent to the 2-digit level. In this context, a new process is defined as a technical improvement in a firm's own production methods, while a new product is an improvement sold to other business enterprises or consumers. Scherer uses data from the Federal Trade Commission's Line of Business survey for 1974 to construct a match between industrial invention patents and the underlying R&D expenditures. He also derives, for each patent, its industry of origin and industries using the invention. Based on these data, Scherer implements a methodology first proposed by Schmookler (1966): Constructing a matrix similar to an input-output table, with industries performing R&D and originating inventions comprising the rows, and industries (including end consumers) using those inventions comprising the columns. Each element in the matrix represents the flow of technology from an originating industry to a using one. Diagonal elements indicate process technology. I use this table to derive, for each industry, its share of R&D spent for product innovation, π_i^{prod}, as the sum of off-diagonal elements divided by total R&D expenditures (row-sum).

Additional control variables

The capital measure in efficiency units used by Krusell et al. (2000) is only available at the aggregate U.S. level. Thus, I use the 4-digit SIC figures from the Manufacturing Industry Database for real capital equipment and structures.[71] The National Science Foundation (NSF) provides company and other (except Federal) research and development (R&D) expenditures as a percentage of sales by industry. This R&D proportion is commonly referred to as R&D intensity.[72] The NSF data cover 24 industries that I match to the 358 industries of my sample.[73] The weighted mean of R&D intensity for my sample increases from 2.12 percent in 1963 to 3.28 percent in 1992.

[71] See Bartelsman and Gray (1996) for a documentation of these data and the corresponding investment deflators.

[72] See, for example, Autor et al. (1998), who work with the same NSF data as used here. Machin and Van Reenen (1998) use R&D intensity in an industry-level panel for several OECD countries and report substantial positive effects on the growth of high-skill employment and wage-bill shares.

[73] The corresponding crosswalk from the 24 NSF industries to the 358 SIC industries of my sample is available upon request. Due to missing observations in the NSF data, several imputations and interpolations were required.

In order to control for computer equipment and other high-technology capital, I use detailed data on private nonresidential fixed assets from the BEA. These data distinguish capital by asset type for 21 (approximately two-digit) NAICS manufacturing industries, which I match to the 358 industries of my panel. I derive the real net capital stock by asset type and industry (in 2000 dollars) from the current-cost capital stock and the chain-type quantity index. Following Berndt, Morrison, and Rosenblum (1992), who use an earlier version of this dataset, I define high-technology capital to include office, computing and accounting machinery; communications equipment; scientific and engineering instruments; and photocopy and related equipment. From this number I calculate the share of high-technology capital in the total capital stock for each industry (*HT/K*). The weighted average of this broad measure increases from 1.2 percent in 1967 to 3.2 percent in 1982, and 6.0 percent in 1992. A frequently used, more narrowly defined measure includes only the share of office, computing and accounting equipment in the capital stock (*OCAM/K*). This variable is 0.4 percent in 1967, 0.8 percent in 1982, and then increases to 2.0 percent in 1992.

Feenstra and Hanson (1999) derive, for each 4-digit SIC industry, a proxy for imported intermediate inputs from trade data. Expressing this measure relative to total expenditure on non-energy intermediates in each industry gives their broad measure of foreign outsourcing. The narrow measure considers only inputs that are purchased from the same 2-digit SIC industry as the good being produced. While the broad measure includes all imported intermediates, the narrow measure restricts attention to the outsourcing of production activities that could have been performed by the respective industry within the United States. I calculate both measures of outsourcing for the years and sectors included in my sample, using data on U.S. imports and exports by 4-digit SIC industries from the Center for International Data at UC Davis together with the above described input-output data.[74] The weighted averaged broad (narrow) measure increases from 4.4 (2.4) percent in 1967 to 8.6 (3.9) percent in 1982 and 13.4 (6.6) percent in 1992.

3.7.2 Appendix A.2 Equilibrium for the Symmetric Case

Firms' optimization with respect to H_i and L_i yields the relative demand for skilled workers, shown in (3.16). The first order condition (FOC) for $\overline{x}_i$ gives sector i's demand for effective units of each intermediate input j, x_{ij}, as a function of total output and goods prices:

$$x_{ij} = \overline{x}_i = \frac{(1-\alpha)p_i Y_i}{\sum_{j\neq i} p_j a_{ij}},\ \forall j \tag{A.1}$$

In the following, I use these FOC to derive the demand for each factor and the marginal cost of production, which equals the product price under perfect competition. Rearranging (3.16) and substituting for H_i in (3.10) yields

$$L_i = \left(\frac{w_L}{1-\gamma_i}\right)^{-\epsilon} \Omega_i^{\epsilon}\ (\overline{x}_i)^{-\frac{1-\alpha}{\alpha}} \left(\frac{Y_i}{A_i}\right)^{\frac{1}{\alpha}} \tag{A.2}$$

and similarly for H_i:

$$H_i = \left(\frac{w_H}{\gamma_i}\right)^{-\epsilon} \left(e^{\phi_i \sigma_i}\right)^{\epsilon-1} \Omega_i^{\epsilon}\ (\overline{x}_i)^{-\frac{1-\alpha}{\alpha}} \left(\frac{Y_i}{A_i}\right)^{\frac{1}{\alpha}} \tag{A.3}$$

[74]I construct a crosswalk to match the 450 manufacturing industries from the trade database to the 358 industries of my sample. The correspondences are available upon request. Feenstra and Hanson use nominal input shares when calculating the outsourcing measure. My results are robust to using both nominal and real input shares.

where Ω_i is the cost of the H_i-L_i labor composite, given by

$$\Omega_i = \left[\gamma_i^{\epsilon} w_H^{1-\epsilon} \left(e^{\phi_i \sigma_i}\right)^{\epsilon-1} + (1-\gamma_i)^{\epsilon} w_L^{1-\epsilon}\right]^{\frac{1}{1-\epsilon}} \tag{A.4}$$

The next steps lead to factor demand as linear functions of Y_i. Multiplying (A.2) and (A.3) by the respective wages and adding up yields the total cost of labor in sector i:

$$w_L L_i + w_H H_i = \Omega_i \left(\overline{x}_i\right)^{-\frac{1-\alpha}{\alpha}} \left(\frac{Y_i}{A_i}\right)^{\frac{1}{\alpha}} \tag{A.5}$$

The FOC of producers' optimization also yield the standard result that the expenditure share for labor is α, i.e., $w_L L_i + w_H H_i = \alpha p_i Y_i$. Plugging this into (A.1) gives

$$\frac{w_L L_i + w_H H_i}{\overline{p}_i \overline{x}_i} = \frac{\alpha}{(1-\alpha)} \tag{A.6}$$

where $\overline{p}_i \equiv \sum_{j \neq i} a_{ij} p_j$ is the effective (or weighted average) input price. Plugging (A.6) into (A.5) yields the demand for effective units of each input as a function of factor prices and output:

$$\overline{x}_i = \frac{1}{A_i} \frac{1-\alpha}{\overline{p}_i} \left(\frac{\overline{p}_i}{1-\alpha}\right)^{1-\alpha} \left(\frac{\Omega_i}{\alpha}\right)^{\alpha} Y_i \tag{A.7}$$

Using this result together with (A.2) gives the demand for low-skilled labor L_i; and together with (A.3) for high-skilled labor H_i, as functions of factor prices and output:

$$L_i = \frac{1}{A_i} \alpha \left(\frac{1-\gamma_i}{w_L}\right)^{\epsilon} \Omega_i^{\epsilon-1} \left(\frac{\overline{p}_i}{1-\alpha}\right)^{1-\alpha} \left(\frac{\Omega_i}{\alpha}\right)^{\alpha} Y_i \tag{A.8}$$

$$H_i = \frac{1}{A_i} \alpha \left(\frac{\gamma_i}{w_H}\right)^{\epsilon} \left(e^{\phi_i \sigma_i}\right)^{\epsilon-1} \Omega_i^{\epsilon-1} \left(\frac{\overline{p}_i}{1-\alpha}\right)^{1-\alpha} \left(\frac{\Omega_i}{\alpha}\right)^{\alpha} Y_i \tag{A.9}$$

We can now derive the total cost of production, TC_i, by multiplying (A.7) - (A.9) with the corresponding factor prices and adding up.[75]

$$TC_i = \frac{1}{A_i} \left(\frac{\overline{p}_i}{1-\alpha}\right)^{1-\alpha} \left(\frac{\Omega_i}{\alpha}\right)^{\alpha} Y_i \tag{A.10}$$

Due to perfect competition within sectors and constant returns to scale in production, representative firms make zero profits, implying $p_i Y_i = TC_i$. Therefore, the price of good i is given by

$$p_i = \frac{1}{A_i} \left(\frac{\overline{p}_i}{1-\alpha}\right)^{1-\alpha} \left(\frac{\Omega_i}{\alpha}\right)^{\alpha}. \tag{A.11}$$

We can now derive the quantities for the symmetric case described in Definition 1. First, from (3.13): $\sigma_i = 1/(N-1)\sum_{j \neq i} h_j$; the input skill intensity of sector i is equal to the average skill intensity of production in all other sectors. Plugging this result into (3.16) and using $H_i/L_i = h_i/(1-h_i)$ gives:

$$\frac{h_i}{1-h_i} = \left(\frac{\gamma}{1-\gamma}\right)^{\epsilon} \left(e^{\phi \sigma_i}\right)^{\epsilon-1} \left(\frac{w_L}{w_H}\right)^{\epsilon} \tag{A.12}$$

[75] Recall that $\overline{x}_i$ reflects also the total amount of inputs used in sector i, which follows from (3.12) and the normalization $\sum_{j \neq i} a_{ij} = 1$. The total cost of intermediate inputs is equal to $\sum_{j \neq i} p_j X_{ij} = \sum_{j \neq i} p_j a_{ij} \overline{x}_i = \overline{p}_i \overline{x}_i$, i.e., weighted average input price times total amount of inputs used.

This equation implies that $h_i = h = H/(H+L), \forall i$.[76] Plugging this into (A.4) yields $\Omega_i = \Omega, \forall i$. Moreover, the input skill intensity is equal to the average high-skill labor share in each sector: $\sigma_i = h, \forall i$. Next, using $\overline{p}_i = 1/(N-1)\sum_{j\neq i} p_j$ in (A.11) implies $p_i = p, \forall i$.[77] Consequently, $\overline{p}_i = p, \forall i$. Because of price symmetry, final demand (3.17) is also symmetric, and so are factor demands (A.7)-(A.9). Thus, $L_i = L/N$, $H_i = H/N$, and $Y_i = Y/N$, where Y is total (intermediate and final) output of the economy. Dividing (A.9) by (A.8) in the symmetric case gives equation (3.21).

Finally, I show that goods markets clear, using the superscripts D for demand and S for supply. Total demand for each good i has a final and an intermediate component: $Y_i^D = C_i + X_{\bullet i}$. The former derives from (3.17) and is given by

$$C_i = c_{L,i}L + c_{H,i}H = \frac{w_L L + w_H H}{Np}, \tag{A.13}$$

while the latter is composed of the demand for sector i's output from all other $N-1$ sectors:

$$X_{\bullet i} = \sum_{j\neq i} \frac{1}{N-1}\overline{x}_j = \frac{1}{A}\left(\frac{p}{1-\alpha}\right)^{1-\alpha}\left(\frac{\Omega}{\alpha}\right)^{\alpha}\frac{1-\alpha}{p(N-1)}\sum_{j\neq i} Y_j^S = \frac{(1-\alpha)}{N-1}\sum_{j\neq i} Y_j^S \tag{A.14}$$

where the first equality follows from (3.12), and the last one from (A.11). In order to join these two equations, I replace Y_j^S using the symmetric expression for the labor expenditure share, $\alpha p Y_j^S = (w_L L + w_H H)/N$ for all sectors $j \neq i$. Therefore, total demand for each good i is given by

$$Y_i^D = C_i + X_{\bullet i} = \frac{w_L L + w_H H}{Np} + \frac{(1-\alpha)}{\alpha(N-1)}\sum_{j\neq i}\frac{w_L L + w_H H}{Np} = \frac{1}{\alpha}\frac{w_L L + w_H H}{Np}. \tag{A.15}$$

The total demand for i is therefore a multiple $1/\alpha$ of the corresponding final demand. With $\alpha = 0.5$, doubling final demand means quadrupling total demand. Under perfect competition, in each sector i total sales equal total expenditures for labor and intermediates:

$$p_i Y_i^S = w_L L_i + w_H H_i + \sum_{j\neq i} p_j X_{ij} \tag{A.16}$$

where the last term is equal to $\overline{p}_i \overline{x}_i$. Under symmetry, and using (A.7) together with the labor expenditure share to replace $\overline{p}_i \overline{x}_i$, (A.16) yields total supply for each i

$$Y_i^S = \frac{1}{\alpha}\frac{w_L L + w_H H}{Np}, \tag{A.17}$$

which equals total demand given by (A.15).

3.7.3 Appendix A.3 The Multiplier in the Model with N Sectors

Suppose that an exogenous innovation arrives in sector i, augmenting skill demand by δ_i. The total change in sector i's high-skilled labor share is then given by $\triangle h_i^T = \triangle h_i^E + \delta_i$, where $\triangle h_i^E$ denotes

[76] To prove this result, note that $h = [(N-1)/N]\sigma_i + [1/N]h_i$. Now suppose that sector i uses more than the average skilled labor share, $h_i > h$. Then $\sigma_i < h$. However, (A.12) requires that $\sigma_i > h$ in order to have $h_i > h$. A similar contradiction arises when we suppose $h_i < h$.

[77] The proof is similar to the one in the previous footnote. Define the average price as $p = [(N-1)/N]\overline{p}_i + [1/N]p_i$ and suppose that sector i charges more, $p_i > p$. Then $\overline{p}_i < p$, which implies that sector i has a lower intermediate input price, but charges more for its final product than the economy average, therefore making positive profits. This would attract competitors charging lower prices until $p_i = p$.

the endogenous component due to the multiplier effect. The latter is driven by changes in i's input skill intensity:

$$\triangle h_i^E = \beta \sum_{j \neq i} a_{ij} \triangle h_j^T \tag{A.18}$$

where the sum corresponds to $\triangle \sigma_i$, and β is the strength of the ITSC, as estimated in section 3.4. Now suppose symmetry, such that $h_i = h_j = h$, $\delta_i = \delta_j = \delta$ and $a_{ij} = 1/(N-1)$. Then (A.18) simplifies to $\triangle h^E = \beta(\triangle h^E + \delta)$, which implies

$$\triangle h^E = \frac{\beta}{1-\beta}\delta \quad \text{and} \quad \triangle h^T = \frac{1}{1-\beta}\delta. \tag{A.19}$$

Therefore, an exogenous innovation that leads to economy-wide skill-upgrading of 1 percent increases the skilled labor share by $1/(1-\beta)$ percent because of the ITSC. With $\beta \approx .33 - .5$, the multiplier effect augments initial skill upgrading by 50 to 100 percent.

References

[1] Acemoglu, Daron (1998). "Why Do New Technologies Complement Skills? Directed Technical Change And Wage Inequality." Quarterly Journal of Economics 113(4): 1055-1089.

[2] Acemoglu, Daron (2002). "Directed Technical Change." The Review of Economic Studies 69(4): 781-809.

[3] Acemoglu, Daron (2007). "Equilibrium Bias of Technology." Econometrica 75(5): 1371-1409.

[4] Acemoglu, Daron and Fabrizio Zilibotti (2001). "Productivity Differences." Quarterly Journal of Economics 116(2): 563-606.

[5] Angrist, Joshua D. (1995). "The Economic Returns to Schooling in the West Bank and Gaza Strip." American Economic Review 85(5): 1065-87.

[6] Arellano, Manuel, and Stephen R. Bond (1991). "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations." Review of Economic Studies 58(2): 277-297.

[7] Autor, David H.; Lawrence F. Katz and Melissa S. Kearney (2008). "Trends in U.S. Wage Inequality: Revising the Revisionists." Review of Economics and Statistics, forthcoming.

[8] Autor, David H.; Lawrence F. Katz and Alan B. Krueger (1998). "Computing Inequality: Have Computers Changed the Labor Market?" Quarterly Journal of Economics 113 (4): 1169-1214.

[9] Autor, David H.; Frank Levy, and Richard J. Murnane (2003). "The Skill Content of Recent Technological Change: An Empirical Exploration." Quarterly Journal of Economics 118(4): 1279-1333.

[10] Bartel, Ann P., and Frank Lichtenberg (1987). "The Comparative Advantage of Educated Workers in Implementing New Technology." Review of Economics and Statistics 69(1): 1-11.

[11] Bartelsman, Eric J., and Wayne Grey (1996). "The NBER Manufacturing Productivity Database." NBER Technical Working Paper 205.

[12] Basu, Susanto (1995). "Intermediate Goods and Business Cycles: Implications for Productivity and Welfare." American Economic Review 85(3): 512-531.

[13] Becker, Randy, John Haltiwanger, Ron Jarmin, Shawn Klimek, and Dan Wilson (2006). "Micro and Macro Data Integration: The Case of Capital." In: A New Architecture for the U.S. Nation Accounts (eds. Jorgenson, Landefeld, and Nordhaus), University of Chicago Press.

[14] Berman, Eli; John Bound, and Zvi Griliches (1994). "Changes in the Demand for Skilled Labor within U.S. Manufacturing: Evidence from the Annual Survey of Manufacturers." Quarterly Journal of Economics 109(2): 367-397.

[15] Berman, Eli; John Bound, and Stephen Machin (1998). "Implications of Skill-Biased Technological Change: International Evidence." Quarterly Journal of Economics, 113(4): 1245-1279.

[16] Berndt, Ernst R., Catherine J. Morrison, and Larry S. Rosenblum (1992). "High-Tech Capital Formation and Labor Composition in U.S. Manufacturing Industries: An Exploratory Analysis." NBER Working Paper No. 4010.

[17] Bertrand, Marianne; Esther Duflo, and Sendhil Mullainathan (2004). "How Much Should We Trust Differences-in-Differences Estimates?" Quarterly Journal of Economics 119(1): 249-275.

[18] Blalock, Garrick, and Paul Gertler (2003). "Technology from Foreign Direct Investment and Welfare Gains through the Supply Chain." Working paper, Haas School of Business, UC Berkeley.

[19] Blanchard, Olivier, and Michael Kremer (1997). "Disorganization." Quarterly Journal of Economics 112(4): 1091-1126.

[20] Bound, John, and George Johnson (1992). "Changes in the Structure of Wages in the 1980s: An Evaluation of Alternative Explanations." American Economic Review 82(3): 371-392.

[21] Bresnahan, Timothy F.; Erik Brynjolfsson, and Lorin M. Hitt (2002). "Information Technology, Workplace Organization and the Demand for Skilled Labor: Firm-Level Evidence." Quarterly Journal of Economics 117(1): 339-376.

[22] Card, David and John E. DiNardo (2002). "Skill-Biased Technological Change and Rising Wage Inequality: Some Problems and Puzzles." Journal of Labor Economics 20(4): 733-783.

[23] Ciccone, Antonio (2002). "Input Chains and Industrialization." Review of Economic Studies 69(3): 565-587.

[24] Ciccone, Antonio, and Giovanni Peri (2005). "Long-run Substitutability Between More and Less Educated Workers: Evidence from U.S. States 1950-1990." Review of Economics and Statistics 87(4): 652-663.

[25] Cohen, Wesley M., and Richard C. Levin (1989). "Empirical Studies of Innovation and Market Structure." R. Schmalensee and R. Willig (ed.) Handbook of Industrial Organization, chapter 18: 1059-1107. Elsevier.

[26] DiNardo, John E. and Jorn-Steffen Pischke (1997). "The Returns to Computer Use Revisited: Have Pencils Changed the Wage Structure Too?" Quarterly Journal of Economics 112(1): 291-303.

[27] Doms, Mark, Timothy Dunne, and Kenneth R. Troske (1997). "Workers, Wages, and Technology." Quarterly Journal of Economics 112(1): 253-290.

[28] Epifani, Paolo and Gino Gancia (2006). "Increasing Returns, Imperfect Competition, and Factor Prices." Review of Economics and Statistics 88(4): 583-598.

[29] Feenstra, Robert C. (2004). "Advanced International Trade: Theory and Evidence." Princeton, N.J. Princeton University Press cop.

[30] Feenstra, Robert C., and Gordon H. Hanson (1999). "The Impact of Outsourcing and High-Technology Capital on Wages: Estimates for the United States, 1979Ű1990." Quarterly Journal of Economics 114(3): 907-940.

[31] Goldin, Claudia, and Lawrence F. Katz (1998). "The Origins of Technology-Skill Complementarity." Quarterly Journal of Economics 113(3): 693-732.

[32] Griliches, Zvi (1969). "Capital-Skill Complementarity." Review of Economics and Statistics 51: 465-468.

[33] Hirschman, Albert O. (1958). The Strategy of Economic Development. New Haven, CT: Yale University Press.

[34] Jones, Charles I. (2007). "Intermediate Goods and Weak Links: A Theory of Economic Development." U.C. Berkeley Working Paper.

[35] Katz, Lawrence F., and David H. Autor (1999). "Changes in the Wage Structure and Earnings Inequality." In Handbook of Labor Economics, vol. 3, O. Ashenfelter and D. Card, eds., (Amsterdam, North Holland: El Sevier).

[36] Katz, Lawrence F., and Kevin M. Murphy (1992). "Changes in Relative Wages, 1963-1987: Supply and Demand Factors." Quarterly Journal of Economics 107(1): 35-78.

[37] Keller, Wolfgang (2004). "International Technology Diffusion." Journal of Economic Literature 42(3): 752-782.

[38] Kim, H. Youn (1992). "The Translog Production Function and Variable Returns to Scale." Review of Economics and Statistics 74 (3): 546-552.

[39] Koo, Jun (2005). "Technology Spillovers, Agglomeration, and Regional Economic Development." Journal of Planning Literature 20(2): 99-115.

[40] Kremer, Michael (1993). "The O-Ring Theory of Economic Development." Quarterly Journal of Economics 108(3): 551-575.

[41] Krueger, Alan B. (1993). "How Computers Changed the Wage Structure: Evidence from Microdata, 1984-1989." Quarterly Journal of Economics 108(1): 33-60.

[42] Krusell, Per; Lee E. Ohanian; José-Víctor Ríus-Rull, and Giovanni L. Violante (2000). "Capital-Skill Complementarity and Inequality: A Macroeconomics Analysis." Econometrica 68(5): 1029-1053.

[43] Leamer, Edward E. (1996). "Wage Inequality from International Competition and Technological Change: Theory and Country Experience." American Economic Review 86(2): 309-314.

[44] Leontief, Wassily (1936). "Quantitative Input and Output Relations in the Economic System of the United States." Review of Economics and Statistics 18(3): 105-125.

[45] Levy, Frank, and Richard J. Murnane (1996). "With What Skills are Computers a Complement?" American Economic Review Papers and Proceedings, LXXXVI (1996): 258-262.

[46] Machin, Stephen, and John Van Reenen (1998). "Technology and Changes in Skill Structure: Evidence from Seven OECD Countries." Quarterly Journal of Economics 113(4): 1215-1244.

[47] Marshall, Alfred (1927). "Principles of Economics." London: Macmillan.

[48] Milgrom, Paul, and John Roberts (1990). "The Economics of Modern Manufacturing: Technology, Strategy, and Organization." American Economic Review 80(3): 511-528.

[49] Nelson, Richard R., and Sidney Winter (1977). "In Search of Useful Theory of Innovation." Research Policy 6(1): 36-76.

[50] Nelson, Richard R., and Edmund S. Phelps (1966). "Investment in Humans, Technological Diffusion, and Economic Growth." American Economic Review 56(1/2): 69-75.

[51] Ngai, L. Rachel and Roberto M. Samaniego (2007). "An R&D-based Model of Multi-sector Growth." London School of Economics Working Paper.

[52] Nunn, Nathan (2007). "Relationship-Specificity, Incomplete Contracts and the Pattern of Trade." Quarterly Journal of Economics 122 (2):569-600.

[53] Pavcnik, Nina (2003). "What explains skill upgrading in less developed countries?" Journal of Development Economics 71: 311-328.

[54] Pavitt, Keith (1984). "Patterns of Technical Change: Towards a Taxonomy and a Theory." Research Policy 13(6): 342-373.

[55] Rauch, James E. (1999). "Networks versus Markets in International Trade." Journal of International Economics 48: 7-35.

[56] Robbins, Carol A. (2006). "The Impact of Gravity-Weighted Knowledge Spillovers on Productivity in Manufacturing." Journal of Technology Transfer 31: 45-60.

[57] Rodríguez-Clare, Andrés (1996). "Multinationals, Linkages, and Economic Development." American Economic Review 86(4): 852-873.

[58] Sanders, Mark, and Bas Ter Weel (2000). "Skill-Biased Technical Change: Theoretical Concepts, Empirical Problems and a Survey of the Evidence." Working paper.

[59] Scherer, Frederick M. (1982). "Inter-industry Technology Flows in the United States." Research Policy 11(4): 227-245.

[60] Schmookler, Jacob (1966). "Invention and Economic Growth." Cambridge: Harvard University Press.

[61] Staiger, Douglas and James H. Stock (1997). "Instrumental Variables Regression with Weak Instruments." Econometrica 65(3): 557-586.

[62] Stock, James H. and Motohiro Yogo (2002). "Testing for Weak Instruments in Linear IV Regressions." NBER Technical Working Paper 284.

[63] Terrell, Katherine (1992). "Productivity of Western and Domestic Capital in Polish Industry." Journal of Comparative Economics 16: 494-514.

[64] Voigtländer, Nico (2008). "Climbing Intermediate Rungs: The Quality Ladder to Wage Inequality." Work in progress.

[65] Wood, Adrian (1998). "Globalisation and the Rise in Labour Market Inequalities." Economic Journal 108(450): 1463-1482.

[66] Wooldridge, Jeffrey M. (2002). "Econometric Analysis of Cross Section and Panel Data." Cambridge, Mass. MIT Press.

[67] Yi, Kei-Mu (2003). "Can Vertical Specialization Explain the Growth of World Trade?" Journal of Political Economy 111(1): 52-102.

[68] Zhu, Susan Chun (2005). "Can product cycles explain skill upgrading?" Journal of International Economics 66: 131Ű155.