NBER WORKING PAPER SERIES

NEOCLASSICAL GROWTH AND THE ADOPTION OF TECHNOLOGIES

Diego Comin
Bart Hobijn

Working Paper 10733
http://www.nber.org/papers/w10733

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
August 2004

The views expressed in this paper solely reflect those of the authors and not necessarily those of the Federal Reserve Bank of New York, nor those of the Federal Reserve System as a whole. We would like to thank Bess Rabin and Rebecca Sela for their great research assistance. We have benefited a lot from comments and suggestions by Jess Benhabib, Simon Gilchrist, Boyan Jovanovic, John Leahy, and Peter Rousseau, as well as seminar participants at NYU, the SED, and UC Santa Cruz.The views expressed herein are those of the author(s) and not necessarily those of the National Bureau of Economic Research.

Neoclassical Growth and the Adoption of Technologies
Diego Comin and Bart Hobijn
NBER Working Paper No. 10733
August 2004
JEL No. E13, O14, O33, O41

ABSTRACT

We introduce a growth model of technology diffusion and endogenous Total Factor Productivity (TFP) levels both at the sector and aggregate level. At the aggregate, the model behaves as the Neoclassical growth model. Our goal is for this model to bridge the gap between the theoretical and empirical studies of technology adoption and economic growth. We bridge this gap in three important directions.

First of all, we use our model to show how one unified theoretical framework is broadly consistent with the observed dynamics of both economic growth as well as of many different measures of technology adoption, like adoption rates, capital to output ratios, and output ratios. Secondly, we estimate our model using a broad range of technological adoption measures, covering 17 technologies and 21 industrialized countries over the past 180 years. This allows us to show how its predicted adoption patterns fit those observed in the data. Finally, we estimate the disparities in sectoral productivity levels as well as aggregate TFP that can be attributed to the differences in the range of technologies in use across countries.

These disparities are almost completely determined by the quality of the worst technology in use, rather than by the quality of the newest technology that has just been adopted or by the number of technologies in use. Further, we find that the TFP component attributable to the range of technologies used is highly correlated with overall sectoral TFP differences across countries, though the variance is smaller.

Diego Comin
Department of Economics
New York University
269 Mercer Street, 725
New York, NY 10003
and NBER
diego.comin@nyu.edu

Bart Hobijn
Domestic Research Function
Federal Reserve Bank of New York
33 Liberty Street
3rd Floor
New York, NY 10045
bart.hobijn@ny.frb.org

1. Introduction

Recent evidence, like Jerzmanowski (2002), Hall and Jones (1999) and Klenow and Rodríquez-Clare (1997), suggests that the bulk of cross-country differences in levels of output per capita is due to differences in the level of total factor productivity (TFP) rather than differences in the levels of factor inputs. These cross-country TFP disparities can essentially be divided into two parts. The first part consists of the productivity differences due to countries using different ranges of technologies. The second part is due to the different levels of efficiency with which all technologies are operated.

There are several theories that propose mechanisms that might lead countries to adopt different sets of available technologies. Among these are Basu and Weil (1998) and Parente and Prescott (1994) for example. The problem with theories that claim that technology adoption lags, i.e. the first component of TFP disparities described above, are an important factor part of TFP differences is that there is little or no empirical evidence to this extend. There are two important reasons why the empirical growth literature trails the theoretical one in this respect.

First and foremost there used to be no data sets that would allow for an extensive analysis of cross-country differences in the adoption patterns of several technologies. Because of this lack of data, most of the studies on technology adoption differences focused on one particular technology, a small number of countries, and only a few points in time. Recent examples of this approach are Saxonhouse and Wright (2000) and Caselli and Coleman (2001). However, Comin and Hobijn (2004) introduced a historical cross-country data set on technology adoption that covers not only a relatively broad sample of countries but also a broad range of technologies as well as a very long time period.

The second reason for the gap between the empirical and theoretical literature on technology adoption is that there are not many macroeconomic models that provide a mapping between their theoretical concept of technological progress and variables that are actually measurable in the data. On the theoretical side, most models provide predictions on the path of capital labor and capital output ratios. On the empirical side, conventional measures of technology adoption, like the adoption shares analyzed in Griliches (1957), Mansfield (1961), and Gort and Klepper (1982), are not directly interpretable in the context of these variables. They are also harder to compute, since they require a measure of the number of potential adopters.

In this paper we introduce a growth model where different firms simultaneously produce intermediate goods with various productivity levels. New intermediate goods are potentially more productive, but are initially operated below their full-efficiency. The array of intermediate goods adopted for production is endogenous. At the micro level, our model is a model of slow diffusion of technologies (embobied in new intermediate goods) qualitatively consistent with the stylized facts: S-shaped diffusion curves and investment in old vintages. The endogenous adoption of new technologies provides a theory of TFP both at the sector and aggregate level. Further, the aggregate behavior of our model is identical to the Neoclassical growth model.

The theoretical literature on growth and diffusion is quite thin. Chari and Hopenhayn (1991) present the model that we believe comes closest to ours. There are several important differences between Chari and Hopenhayn (1991) and our model: First, our model has capital. This is important because most of the technologies that we measure are embodied. Chari and Hopenhayn instead model disembodied technologies.

Second, our model aggregates to the Neoclassical growth model. This greatly simplifies the investigation of the aggregate and sectoral behavior of our model along the transition path. But, beyond the independent interest of our model, what is important for our purposes is that it is very well suited to bridge the existing gap between the theory on growth and adoption and the data.

In the tradition of the growth literature, our model provides predictions for the path of aggregate capital-output ratios. It also has implications for the paths of technology specific capital-output and output ratios. It shows how the latter can be used as measures of technology adoption and what growth theory predicts about their dynamic behavior. It also shows how these measures map into the more conventional adoption shares, which were used by Griliches (1957) and Mansfield (1961). The advantage of developing a theory in terms of capital-output and output ratios is that these correspond directly to the cross-country measures of technology adoption for which we have most data.

We use the Historical Cross-Country Adoption Data set (HCCTAD), introduced in Comin and Hobijn (2004), to estimate our model. Specifically, we use data for 17 technologies in 21 industrialized countries spanning 180 years. We show that the dynamics of technology adoption predicted by the model are qualitatively and quantitatively consistent with the general patterns of technology adoption observed in the data for this broad range of technologies and countries. Our results suggest that there are substantial differences in the rate at which countries absorb new technologies and increase the efficiency level at which they use them.

Finally, we use the structural estimates that we obtain from our model to measure the productivity differentials between countries that can be attributed to the disparities in the technology adoption patterns that we observe. This component of productivity is determined mainly by the level of the least advanced technology in use. We call this margin capital dragging. We observe also that different countries tend to lead in the productivity component associated with technology adoption for different technologies.

These differences are, in general, significantly correlated with overall sectoral productivity disparities. However, their variance is substantially smaller. Hence, our analysis still leaves a large part of sectoral productivity differences to be explained by conventional TFP factors that affect all technologies that are used in the same way.

The structure of this paper is as follows. In the next section, we introduce our model. In Section 3 we derive its equilibrium dynamics. Section 4 contains the main theoretical contribution of the paper. It explains what we assume for the path of vintage specific technological progress in this model and what it implies for the adoption of technologies. Section 5 is devoted to mapping our theoretical concept of technological change into variables that are measured in the data. Section 6 contains a description of our estimation strategy, while Section 7 contains the results obtained by applying it to the data from Comin and Hobijn (2004). Section 8 concludes.

2. Model

Our aim is to model the interaction between economic growth and the adoption of several major technologies. This is why we consider a multi-sector growth model in which each sector uses a major technology.

Households

Households derive utility from a countably finite number of different final goods and services. These goods and services are indexed by $i=1,\ldots,N$. Consumption of each of these goods in period t is denoted by C_{it}. Overall utility is derived from a CES composite of these individual goods. This composite equals

$$C_t = \left[\sum_{i=1}^{N} C_{it}^{\rho}\right]^{1/\rho} \text{ where } 0<\rho<1 \tag{1}$$

The household maximizes the present discounted value of utility derived from its path of consumption $\{C_{t+s}\}_{s=0}^{\infty}$. This present discounted value is given by

$$\frac{\sigma}{\sigma-1}\int_0^{\infty} e^{-\beta s} C_{t+s}^{\frac{\sigma-1}{\sigma}}\, ds \tag{2}$$

and is maximized subject to the flow budget constraint

$$\dot{B}_{t+s} = r_{t+s}B_{t+s} + W_{t+s}L_{t+s} + \Pi_{t+s} + \Xi_{t+s} - C_{t+s} \tag{3}$$

Here, B_{t+s} reflects the household's bondholdings, r_{t+s} is the real interest rate at time $t+s$, W_{t+s} is the real wage rate, Π_{t+s} are the aggregate real flow profits that are distributed to the households, and Ξ_{t+s} represents additional payments that households receive from businesses. We abstract from employment fluctuations and assume that the representative household inelastically supplies a unit measure of labor, such that $L_{t+s}=1$ for all $s=[0,\infty)$.

Note that we have written the budget constraint such that the consumption composite C_t is the numeraire good, which by definition has a price equal to one in every period. Let P_{it} be the price of good i at time t and let $P_t=1$ be the price of the consumption composite C_t. The CES structure that we have assumed implies that in that case

$$P_t = \left[\sum_{i=1}^{N}\left(\frac{1}{P_{it}}\right)^{\frac{\rho}{1-\rho}}\right]^{-\frac{1-\rho}{\rho}} = 1 \tag{4}$$

In terms of the consumption aggregate, this optimization problem yields an Euler equation of the form

$$\frac{\dot{C}_t}{C_t} = \sigma\left[r_t - \beta\right] \tag{5}$$

Demand for the individual goods is implied by the CES structure and determined by

$$C_{it} = P_{it}^{-\frac{1`}{1-\rho}} C_t \tag{6}$$

This is the consumption demand function that the final goods producers face.

Final goods sectors

This economy has N final goods sectors, each of which produces a different final good or service indexed by i. The sector that produces the final good of type i consists of a continuum of perfectly competitive firms that each use a constant returns to scale CES production technology.

This allows us to represent the aggregate actions of these firms as if taken by a representative firm. The final good of type i is produced using a continuum of intermediate goods. The continuum of intermediate goods that is used for the production of final good i is denoted by the interval $\left[\underline{v}_{it}, \bar{v}_{it}\right]$. These intermediate goods are combined to produce final good i, using the production function[1]

$$Y_{it} = \left[\int_{\underline{v}_t}^{\bar{v}_t} Y_{ivt}^{\theta} dv\right]^{1/\theta} \quad \text{where } 0<\theta<1 \tag{7}$$

The final good producers sell their output at price P_{it} and buy the intermediate good of type v at price P_{ivt}. This results in the demand functions for the intermediate inputs that read

$$Y_{ivt} = \left(\frac{P_{ivt}}{P_{it}}\right)^{-\frac{1}{1-\theta}} Y_{it} \tag{8}$$

Where P_{it} is determined by the zero profit condition for this perfectly competitive sector and equals

$$P_{it} = \left[\int_{\underline{v}_{it}}^{\bar{v}_{it}} \left(\frac{1}{P_{ivt}}\right)^{\frac{\theta}{1-\theta}} dv\right]^{-\frac{1-\theta}{\theta}} \tag{9}$$

For simplicity, and to allow for aggregation of the various sectors, we will assume that the elasticity of substitution between different intermediate goods, i.e. $1/(1-\theta)$, is constant across the final goods sectors.

Intermediate goods sectors

Each intermediate good used in the production of the final good i is produced using capital goods of a particular technology vintage. In particular, the output of intermediate good v for final good i at time t is assumed to be produced using a Cobb-Douglas production technology of the form

$$Y_{ivt} = Z_{ivt} K_{ivt}^{\alpha} L_{ivt}^{1-\alpha} \tag{10}$$

[1] Levhari (1968) provides a microfoundation of a CES production function in an environment where each producer faces a heterogeneous level of productivity. Closer to our context, Anderson, de Palma, and Thisse (1993) introduced the microfoundations of CES preferences. They show how they can be interpreted as the aggregate utility representation of a market consisting of a continuum of heterogeneous agents each making a discrete choice among the goods available and then choosing how much to consume of that good. This analysis can easily be extended to a production context where firms select which intermediate goods to use.

Here, Z_{ivt} is the vintage specific productivity level, which potentially varies over time, and K_{ivt} is the capital stock that embodies technology vintage v for the production of i. Finally, L_{ivt} is the amount of labor inputs used in the production of Y_{ivt}.

The vintage specific productivity level is similar to that in classic vintage capital models like Johansen (1959) and Solow (1960). What is different here from these models is that Z_{ivt} changes over time.

The intermediate goods producers are monopolistic competitors. The monopoly power that they possess is due to their proprietary knowledge of how to convert a unit of the standardized investment good into a unit of the capital good that embodies technology v for the production of good i. This conversion is reversible so that it also yields a unit of the standard investment good for every unit of K_{ivt} disinvested.

Firms are assumed to take the input prices, i.e. the investment price and the real wage rate, as given. In each period, the supplier of vintage v for good i makes flow profits equal to its revenue minus its capital and labor expenses as well as minus a fixed cost of doing business in sector i in period t. Mathematically, these flow profits can be written as

$$\Pi_{ivt} = P_{ivt}Y_{ivt} - I_{ivt} - W_t L_{ivt} - \Xi_{it} \tag{11}$$

where Ξ_{it} represents the per period cost of operating in sector i in period t.

Each intermediate goods supplier is assumed to maximize the present discounted value of its flow profits over the period that it is in business. Let $\underline{t}_{iv}$ denote the time in which the supplier enters the market. Similarly, let $\bar{t}_{iv}$ denote the time the supplier exits the market. The supplier's objective is to maximize

$$\int_{\underline{t}_{iv}}^{\bar{t}_{iv}} e^{-\int_0^t r_s ds} \Pi_{ivt} dt \tag{12}$$

subject to the technological constraint, (10), the demand function, (8), as well as the capital accumulation equation

$$\dot{K}_{ivt} = I_{ivt} - \delta K_{ivt} \tag{13}$$

Solving the above dynamic profit maximization of the producer of vintage v for good i yields the following optimality condition for labor demand

$$W_t = (1-\alpha)\theta \frac{P_{ivt}Y_{ivt}}{L_{ivt}} \tag{14}$$

where $(1-\alpha)\theta$ is the labor elasticity of revenue. Similar to Jorgenson (1963), the dynamic optimality condition for investment of this supplier is a form of the user cost equation of capital. It reads

$$(r_t + \delta) = \alpha\theta \frac{P_{ivt}Y_{ivt}}{K_{ivt}} \tag{15}$$

Where $(r_t+\delta)$ is the user cost of capital.

So far, we have solved the problem for the supplier under the assumption that he was operating in the market. Here we determine the conditions under which this is actually the case. That is, we derive what determines $\underline{t}_{iv}$ and $\bar{t}_{iv}$.

The supplier of vintage *v* for good *i* will produce as long as its flow profits, excluding it capital expenditures, exceed the user cost its incurs on its capital stock. That is, it will produce as long as

$$\left[P_{ivt}Y_{ivt} - W_t L_{ivt} - \Xi_{it}\right] \geq \left(r_t + \delta\right) K_{ivt} \tag{16}$$

Substituting in the optimality conditions that we derived above, we obtain that this implies that vintage suppliers are in business as long as their revenue satisfies

$$P_{ivt}Y_{ivt} \geq \frac{\Xi_{it}}{1-\theta} \tag{17}$$

Hence, suppliers that do not generate enough revenue to recoup the operating cost in their sector will exit the market.

Investment goods sector

What remains to be defined is how the standardized investment good is produced. Throughout this paper, our aim is to consider technological adoption in a model that is as similar as possible to the benchmark one good neoclassical growth model. Because of this, we have chosen to specify the investment good producing sector in such a way that the prices of the consumption and investment goods are equal at any point in time.

We do so by assuming that the investment good producing sector is perfectly competitive. Each competitor in this market uses investment demand of each final good *i*, which we will denote by I_{it}, in a production function that this similar to the preferences of the household. That is,

$$I_t = \left[\sum_{i=1}^{N} I_{it}^{\rho}\right]^{1/\rho} \tag{18}$$

where I_t is the total output of the standardized investment good.

This choice of production function and the zero profit condition for this perfectly competitive market yield that the price of the investment good has to equal

$$\left[\sum_{i=1}^{N}\left(\frac{1}{P_{it}}\right)^{\frac{\rho}{1-\rho}}\right]^{-\frac{1-\rho}{\rho}} \tag{19}$$

which equals the consumption price level derived in (4), which, in turn, equals one by assumption.

3. Equilibrium and aggregation

Before we solve for the dynamic equilibrium path of this economy, it is worthwhile to first derive some aggregation results that reduce the set of aggregates relevant for the equilibrium dynamics.

Derivation of aggregates

The aggregate representation of the model economy that we develop is identical to that of the standard one-sector neoclassical growth model. Hence, the aggregates that we derive will coincide with the equilibrium variables of the neoclassical growth model. The important difference with the standard neoclassical growth model is that the endogenous adoption of technologies provides a theory for endogenous TFP both at the sector and aggregate level.

Under the optimality conditions derived above, output in each final goods sector can be represented as being produced using a Cobb-Douglas production function of the form[2]

$$Y_{it} = Z_{it} K_{it}^{\alpha} L_{it}^{1-\alpha} \tag{20}$$

where the labor input, L_{it}, is just the total amount of labor demanded by the producers in the sector. That is,

$$L_{it} = \int_{\underline{v}_{it}}^{\bar{v}_{it}} L_{ivt} dv . \tag{21}$$

Writing the capital input in terms of units of the standardized investment good allows us to aggregate capital across the various vintages as follows:

$$K_{it} = \int_{\underline{v}_{it}}^{\bar{v}_{it}} K_{ivt} dv . \tag{22}$$

The total factor productivity level is given by

$$Z_{it} = \left[\int_{\underline{v}_{it}}^{\bar{v}_{it}} Z_{ivt}^{\frac{\theta}{1-\theta}} dv \right]^{\frac{1-\theta}{\theta}} . \tag{23}$$

The theory of endogenous TFP that we present in this paper can be understood by looking at this expression. Recall that, each intermediate good has a specific total factor productivity level (Z_{ivt}). The TFP level in the sector results from two mechanisms. First, an endogenous measure of intermediate goods (determined by the interval $[\underline{v}_{it}, \bar{v}_{it}]$) are used in production. Hence, the TFP level depends on the total factor productivity embodied in the intermediate goods that enter in production. Second, the CES aggregation of the intermediate goods creates an externality from the variety of intermediate goods used in production. This externality shows up in the TFP measure Z_{it}.

What is important for our analysis is not only that the production functions aggregate to a simple representation, but also that the optimality conditions for each sector can be aggregated. That is, as shown in Appendix A, the optimality conditions, derived in the previous section, aggregate to

$$W_t = (1-\alpha)\theta \frac{P_{it} Y_{it}}{L_{it}} \text{ and } (r_t + \delta) = \alpha\theta \frac{P_{it} Y_{it}}{K_{it}} \tag{26}$$

[2] We leave the details of the derivation of these sectoral production functions for Appendix A.

which implies, together with the derived sectoral production function, that the decentralized allocation of resources in each sector is as if decided upon by a single representative firm which equates marginal revenue to marginal cost.

Similar to the Cobb-Douglas vintage production functions aggregating to a Cobb-Douglas sectoral production function, the Cobb-Douglas sectoral production functions aggregate to an aggregate Cobb-Douglas production function. This production function reads

$$Y_t = Z_t K_t^{\alpha} L_t^{1-\alpha} = C_t + I_t \tag{27}$$

and the corresponding total factor productivity level is defined as

$$Z_t = \left[\sum_{i=1}^{N} Z_{it}^{\frac{\rho}{1-\rho}} \right]^{\frac{1-\rho}{\rho}} \tag{28}$$

while the capital and labor aggregates equal

$$K_t = \sum_{i=1}^{N} K_{it} \text{ and } L_t = \sum_{i=1}^{N} L_{it} = 1 \tag{29}$$

As is shown in Appendix A, the optimality conditions also aggregate to the economy wide level. Just like the sectoral input aggregates, the aggregates K_t and L_t satisfy

$$W_t = (1-\alpha)\theta \frac{Y_t}{L_t} \text{ and } (r_t + \delta) = \alpha\theta \frac{Y_t}{K_t} \tag{30}$$

The decisions of the continuum of different producers that are distributed over N sectors in this economy can be represented as if made a by a representative firm using a Cobb-Douglas technology. Moreover, the aggregate capital stock also satisfies the accumulation equation

$$\dot{K}_t = I_t - \delta K_t \tag{31}$$

where aggregate gross investment is defined as

$$I_t = \sum_{i=1}^{N} I_{it} \text{ and } I_{it} = \int_{\underline{v}_{it}}^{\bar{v}_{it}} I_{ivt} dv \tag{32}$$

Thus, at the aggregate level this economy is equivalent to that of the standard one-sector neoclassical growth model. The only difference is that, because of the monopoly power of the intermediate goods suppliers in each sector, factor payments to capital and labor only make up a fraction θ of nominal output.

We will assume that the flow profits of the intermediate goods suppliers as well as the costs of doing business these suppliers incur flow back to the households, in the form of Π_t and Ξ_t respectively[3]. Here

[3] The results in Caselli and Ventura (2000) imply that aggregate household behavior, given the CRRA preferences in our model, does not depend on how the (lump-sum) payments of Π_t and Ψ_t are distributed across households. Therefore, we do not make any specific assumptions about to which households these payments are made.

$$\Pi_t = \sum_{i=1}^{N} \int_{\underline{v}_{it}}^{\bar{v}_{it}} \Pi_{ivt} dv \text{ and } \Xi_t = \sum_{i=1}^{N} \left[\bar{v}_{it} - \underline{v}_{it} \right] \Xi_{it} \tag{33}$$

These two payments make up the remaining fraction, *(1-θ)*, of nominal output.

Aggregate equilibrium dynamics

The preceding results imply that the dynamics of the aggregates in this model are determined by the following two differential equations. The first is the equilibrium consumption Euler equation, which equals

$$\frac{\dot{C}_t}{C_t} = \sigma \left[\theta \alpha Z_t \left(\frac{1}{K_t} \right)^{1-\alpha} - \delta - \beta \right] \tag{34}$$

The second is the capital accumulation equation

$$\frac{\dot{K}_t}{K_t} = Z_t \left(\frac{1}{K_t} \right)^{1-\alpha} - \delta - \frac{C_t}{K_t} \tag{35}$$

Apart from one exception, this system is identical to the one that determines the equilibrium dynamics of consumption and capital in the standard one sector neoclassical growth model. The exception is the inefficiency due to the wedge that the monopolistic competition between the intermediate goods suppliers drives between the social and private marginal return to capital. This is reflected by the factor θ multiplying the social returns to capital in (34).

Sectoral equilibrium dynamics

Underneath the aggregate dynamics of our economy are the sectoral dynamics. These determine the path of technology adoption in each of the N final goods sectors. We explain the details of the derivation of these equilibrium dynamics in Appendix A. The following ratio's make up the gist of them and are what is relevant for the rest of this paper.

First of all, the sectoral to total output ratio and the vintage to sectoral output ratio are given by

$$\frac{Y_{it}}{Y_t} = \left(\frac{Z_{it}}{Z_t} \right)^{\frac{1}{1-\rho}} \text{ and } \frac{Y_{ivt}}{Y_{it}} = \left(\frac{Z_{ivt}}{Z_{it}} \right)^{\frac{1}{1-\theta}} \tag{36}$$

respectively. Moreover, the resulting prices are given by

$$\frac{P_{it}}{P_t} = P_{it} = \left(\frac{Y_t}{Y_{it}} \right)^{1-\rho} = \left(\frac{Z_t}{Z_{it}} \right) \text{ and } \frac{P_{ivt}}{P_{it}} = \left(\frac{Y_{it}}{Y_{ivt}} \right)^{\frac{1}{1-\theta}} = \left(\frac{Z_{it}}{Z_{ivt}} \right) \tag{37}$$

Thus, this derivation allows us to write the sectoral and vintage specific equilibrium quantities in terms of the underlying levels of total factor productivity.

Contrary to the standard neoclassical growth model, in which total factor productivity evolves exogenously, the sectoral total factor productivity level in this economy is determined endogenously. Equation (23) implies that the sectoral total factor productivity level Z_{it} is a function of the range of technology vintages that are operated in period t. Thus, it depends on $\underline{v}_{it}$ and $\overline{v}_{it}$. These vintages are determined by the entry and exit condition (17), which, using the equilibrium results above, can be written as

$$Z_{ivt} \geq \left(\frac{1}{1-\theta} \frac{\Xi_{it}}{P_{it} Y_{it}} \right)^{\frac{1-\theta}{\theta}} Z_{it} = \underline{Z}_{it} \tag{38}$$

Solving this equation requires making a set of additional assumptions about the paths of Z_{ivt} and Ξ_{it}. These assumptions determine the shape and pattern of technological progress in our model. Since these assumptions are one of the major contributions of this paper and, above all, are crucial to the adoption dynamics in our model, we introduce them in detail in the next, separate, section.

4. *Technological progress and adoption*

In the previous section we derived the aggregate equilibrium dynamics of our model economy and showed that, at the aggregate level, this economy is observationally equivalent to the standard neoclassical growth model. The main focus of this paper is the integration of these aggregate dynamics with a theory of technology adoption of various major technologies in the economy.

In this section we introduce our specification of technological progress and show what it implies for the dynamics of technology adoption in our model economy. We do so in three steps. In the first step, we introduce our assumption about the shape of technological progress in this economy. In the second step, we solve for the implied values of $\underline{v}_{it}$ and $\overline{v}_{it}$ under these assumptions. In the third step, we solve for the theoretical measures of technology adoption that result from this solution and that are of interest for our empirical analysis.

The shape of technological progress

The main difference between our model and the most commonly used vintage capital models is that we allow for the vintage specific productivity level, Z_{ivt}, to vary over time. In fact, this variation will be crucial in what is to follow.

For Z_{ivt}, we will assume that it satisfies

$$Z_{ivt} = Z_{iv^*_{it}t} f_i\left(v - v^*_{it}\right) \tag{39}$$

This specification of the vintage specific productivity profile in itself does not provide us with a solution to our model. However, it provides us with a framework to capture the shape of technological progress that we are after.

In order to capture this shape, we make the following additional assumptions. The first is that

$$Z_{iv^*_{it}t} = a_i e^{g_i v^*_{it}} \tag{40}$$

Without loss of generality, we will also assume that $v^*_{it} = v^*_{i0} + t$. This assumption allows us, in equilibrium, to interpret g_i as the rate of technological change in sector i. For the function F_i we assume that

$$F_i(x) = e^{g_i x} \text{ for } x \leq 0,\ \partial \ln F_i(x)/\partial x < g_i \text{ and } \partial^2 \ln F_i(x)/\partial^2 x < 0 \text{ for } x > 0 \text{, and } \lim_{x\to\infty} F_i(x) = 0 \tag{41}$$

The economic implications of this assumption are easiest discussed in the context of an example. For this purpose, consider Figure 1. This figure plots a representative shape of the logarithm of the vintage technology frontier at a particular point in time.

The dashed line depicts the feasibility frontier for each vintage. This is the maximum productivity level that each vintage will converge to as *t* goes to infinity. All vintages below v_{it}^* are operated at this maximum productivity level and can be interpreted as fully matured technologies.

Vintages that are beyond v_{it}^* are operated below their long run potential. In fact, our assumptions on F_i are such that vintages that are further beyond v_{it}^* are operated at a smaller fraction of their long-run maximum. We assume that, in the limit, this fraction goes to zero when $(v - v_{it}^*)$ goes to infinity.

In principle, our specification of technological progress is rather mechanical and we prefer to interpret it as a reduced form specification rather than a microfounded one. There are several papers, however, that provide some insight in which economic processes might generate such a shape of the technological frontier. First of all, the increasing fraction of maximum efficiency at which the technology is operated can be directly interpreted as a learning curve. That is, evidence on learning by doing, as in Bahk and Gort (1993), Irwin and Klenow (1994), and Jovanovic and Nyarko (1995), would generate a vintage productivity profile that is similar to that in Figure 1.

Learning by doing is not the only mechanism that could be behind this pattern, though. If incumbent firms lobby the political system to essentially tax newer entrepeneurs then one might also observe a pattern like this. An example of a model where incumbents apply resources to lobbying to the detriment of their potential competitors is Holmes and Schmitz (2001).

Fernandes (2003) argues that new technologies might not be adopted because complementary financial innovations might be needed to accommodate the risk associated with their use. Although her argument is not directly applicable here, it does point out the fact that the productivity of newer technology vintages might improve over time because markets as well as complimentary inventions adjust to them.

Entry and exit revisited

Our assumptions about the functional form for the shape of the technological frontier in each sector allow us to write the sectoral total factor productivity level as

$$Z_{it} = Z_{iv_{it}^*t}\left[\int_{\underline{d}_{it}}^{\bar{d}_{it}} F_i(x)^{\frac{\theta}{1-\theta}} dx\right]^{\frac{1-\theta}{\theta}} \text{ where } \underline{d}_{it} = (\underline{v}_{it} - v_{it}^*) < (\bar{v}_{it} - v_{it}^*) = \bar{d}_{it} \tag{42}$$

Because the entry and exit condition (38) holds with equality for both $\underline{v}_{it}$ and $\bar{v}_{it}$, we know that

$$F_i(\underline{d}_{it}) = F_i(\bar{d}_{it}) = \left(\frac{1}{1-\theta}\frac{\Xi_{it}}{P_{it}Y_{it}}\right)\left[\int_{\underline{d}_{it}}^{\bar{d}_{it}} F_i(x)^{\frac{\theta}{1-\theta}} dx\right]^{\frac{1-\theta}{\theta}} \tag{43}$$

In order for us to be able to solve this condition we need to make an assumption about the path of Ξ_{it} over time. We will assume that the operating cost in a sector is proportional to the total revenue in that sector. Mathematically

$$\Xi_{it} \equiv \zeta_i P_{it} Y_{it} \tag{44}$$

When this assumption is applied to the entry and exit condition above, we obtain that in every period $\underline{v}_{it}$ and $\bar{v}_{it}$ are determined by the following two conditions

$$F_i(\underline{d}_{it}) = F_i(\bar{d}_{it}) = \left(\frac{\zeta_i}{1-\theta}\right)\left[\int_{\underline{d}_{it}}^{\bar{d}_{it}} F_i(x)^{\frac{\theta}{1-\theta}} dx\right]^{\frac{1-\theta}{\theta}} \text{ where } \underline{d}_{it} = \left(\underline{v}_{it} - v_{it}^*\right) < \left(\bar{v}_{it} - v_{it}^*\right) = \bar{d}_{it} \tag{45}$$

Note that this system does not depend on time. Thus, in equilibrium, $\underline{d}_{it}$ and $\bar{d}_{it}$ are constant over time. Because of this, we will denote them by $\underline{d}_i$ and $\bar{d}_i$ in the rest of this paper.

The above result also means that the width of the range of vintages used in each sector is constant over time[4] and moves because of the progress in v^*_{it} . In fact, we can use the solution for $\underline{d}_i$ and $\bar{d}_i$ to calculate the sectoral level of total factor productivity, which equals

$$Z_{it} = Z_{iv_{it}^*t}\left[\int_{\underline{d}_i}^{\bar{d}_i} F_i(x)^{\frac{\theta}{1-\theta}} dx\right]^{\frac{1-\theta}{\theta}} = a_i e^{g_i t}\left[\int_{\underline{d}_i}^{\bar{d}_i} F_i(x)^{\frac{\theta}{1-\theta}} dx\right]^{\frac{1-\theta}{\theta}} \tag{46}$$

Since the integral term of this expression is constant over time, the growth rate of Z_{it} equals g_i. Hence, our set of assumptions here implies that sectoral total factor productivity levels grow at a constant rate over time.

Graphically, the underlying process of technological change in each sector boils down to Figure 2. In each period the technological frontier moves parallel to the feasibility frontier. This shift makes the productivity level of the set of implemented vintages grow at the rate at which the productivity of v^*_{it} grows. This causes the sectoral TFP level to grow at the same rate as well.

Measures of technological adoption

Our main aim is to derive the path of technological adoption in each of the sectors in the economy. In order to derive this path, we first have to define how we measure technological adoption. The measures that we are interested in are ones that we subsequently are able to measure in the data. In this subsection we introduce the theoretical counterparts of our empirical measures. The available empirical measures are discussed in Section 6.

There are two main measures of interest that we use in our empirical analysis. The first is the vintage output to GDP ratio. Mathematically, this is defined as Y_{ivt}/Y_t. The second is the vintage capital to GDP ratio, which equals K_{ivt}/Y_t.

These two ratios are different from the measures that are most commonly considered in studies of technological progress. Generally, studies of technological progress, like Griliches (1957), Mansfield (1961) and

[4] The constancy of this range is a result of equation (44). Other assumptions about the path of sector specific costs potentially lead to time variation in this range, mainly because they could imply an interaction between sector i's technology choice and macroeconomic aggregates.

Gort and Klepper (1982), for example, consider adoption shares and note that adoption share curves are commonly S-shaped.

As it turns out, our model also generates such curves. They can be derived from the two ratios considered above. In order to show how our model is consistent with previous empirical work on adoption, we will present our theoretical results in terms of the two ratios that we use in our empirical analysis as well as in terms of the adoption shares.

These adoption shares in our model[5] will be the share of output of vintage *v* in total output of sector *i* and the share of capital of vintage *v* in the capital stock of sector *i*. These equal, respectively,

$$Y_{ivt} \Big/ \int_{\underline{v}_{it}}^{\bar{v}_{it}} Y_{iut}du \text{ and } K_{ivt} \Big/ \int_{\underline{v}_{it}}^{\bar{v}_{it}} K_{iut}du \tag{47}$$

The above ratios, Y_{ivt}/Y_t and K_{ivt}/K_t, are perfectly well defined from a theoretical point of view. However, they are not empirically viable yet, because they require measuring data on a continuum of vintages of technologies. Such data are simply not available. Therefore, in order to make our theory empirically applicable, we need to map the continuum of vintages into a countable number of technology vintages that is observed in the data.

5. From adoption in the model to adoption in the data

Where our model has a continuum of vintages for each sector *i*, we only observe a discrete number of technology vintages in the data. This section explains the way we map this continuum of vintages into the discrete number of vintages that we observe in the data. We will start off this explanation with an example and then proceed to generalize from there. After explaining this mapping, we derive what our theory predicts for the various observed measures of technological adoption.

Example: Merchant vessels

Merchant ships have historically been classified into three different categories: (*i*) sailships, (*ii*) steamships, and (*iii*) motorships. Consider merchant shipping services as one of the sectors, *i*, in the model economy. Then our theoretical setup considers a continuum of different technology vintages that are, have been, or potentially could be used to provide shipping services. What we will assume is that the continuum of merchant shipping technology vintages can be split into three intervals, as depicted in Figure 3. The first interval, $(-\infty, v_{steam}]$, is assumed to be the set of vintages that are all sailships. The second, $(v_{steam}, v_{motor}]$ is assumed to be the set of steamship vintages. Finally, we will assume that (v_{motor}, ∞) consists of motorships. For our empirical application, the cut off vintages, v_{steam} and v_{motor} are both unknown parameters. We do have some historical information, though, that allows us with approximate values for these parameters. The first steamboat was built in 1788 in Philadelphia by Fitch. Hence, v_{steam} probably corresponds to date shortly after 1788. The diesel engine, which propels most motorships, was introduced in 1892. Therefore, we list this as the approximate date for v_{motor} in the figure.

[5] In the empirical adoption literature, these share are generally defined in terms of the fraction of firms rather than the fraction of the capital stock or output.

Mapping into a discrete number of vintages

Our data cover several major technologies, which we group into technology groups, like merchant shipping in the example above. For our empirical application we use these technology groups as the equivalent of our sectors *i*.

Within each group, we order the technologies sequentially; like sailships, steamships and motorships are ordered in the example above. We denote these technologies within each group by the index τ. For each of these observed technologies, we then assume that they correspond to a set of vintages, which we denote by $V_i(\tau)$, such that

$$V_i(\tau) = \begin{cases} (-\infty, v_{i\tau+1}] & \text{for} \quad \text{technologies without a predecessor} \\ (v_{i\tau}, v_{i\tau+1}] & \text{for} \quad \text{technologies with both a predecessor and successor} \\ (v_{i\tau}, \infty) & \text{for} \quad \text{technologies without a successor (yet)} \end{cases} \tag{48}$$

The adoption measures for the continuum of vintages that we considered before can be converted into measures for the countably finite number of vintages that we have data for by integrating them over the above technology sets, $V_i(\tau)$.

Predicted technology adoption patterns

Now that we have defined the theoretical equivalent of the observed technologies in our data set, we can derive what our model predicts about the patterns of adoption of these technologies.

First of all, consider the output and capital to GDP ratios. For technology τ of group *i*, we will denote these ratios by $m_{i,Y,t}(\tau)$ and $m_{i,K,t}(\tau)$ respectively. These ratios equal

$$m_{i,Y,t}(\tau) = \int_{V_i(\tau)} (Y_{ivt}/Y_t)dv \text{ and } m_{i,K,t}(\tau) = \int_{V_i(\tau)} (K_{ivt}/Y_t)dv \tag{49}$$

They map into the adoption shares that have been used in a broad number of studies in the following way

$$s_{i,Y,t}(\tau) = \frac{\int_{V_i(\tau)} Y_{ivt}dv}{\sum_s \int_{V_i(s)} Y_{iut}du} = \frac{m_{i,Y,t}(\tau)}{\sum_s m_{i,Y,t}(s)} \text{ and } s_{i,K,t}(\tau) = \frac{\int_{V_i(\tau)} K_{ivt}dv}{\sum_s \int_{V_i(s)} K_{iut}du} = \frac{m_{i,K,t}(\tau)}{\sum_s m_{i,K,t}(s)} \tag{50}$$

The calculation of these shares in (50) requires having observations on all technology vintages used, while the calculation of the ratios in (49) only requires data on one technology vintage and a measure of real GDP.

In appendix A we derive the path of the adoption measures as a function of time. This yields

$$m_{i,Y,t}(\tau) = \left(\frac{Z_{it}}{Z_t}\right)^{\frac{1}{1-\rho}} \left(\frac{Z_{iv_{it}^*t}}{Z_{it}}\right)^{\frac{1}{1-\theta}} \xi_{i,Y,t}(\tau) \text{ and } m_{i,K,t}(\tau) = \left(\frac{Z_{it}}{Z_t}\right)^{\frac{\rho}{1-\rho}} \left(\frac{Y_t}{Z_t^{\frac{1}{1-\alpha}}}\right)^{\frac{1-\alpha}{\alpha}} \left(\frac{Z_{iv_{it}^*t}}{Z_{it}}\right)^{\frac{\theta}{1-\theta}} \xi_{i,K,t}(\tau) \tag{51}$$

where

$$\xi_{i,Y,t}(\tau)=\left[\int_{\min\{\tilde{v}_{i\tau+1}-t,\max\{\tilde{v}_{i\tau}-t,\underline{d}_i\}\}}^{\max\{\tilde{v}_{i\tau}-t,\min\{\tilde{v}_{i\tau+1}-t,\bar{d}_i\}\}} F_i(x)^{\frac{1}{1-\theta}}dx\right] \text{ and } \xi_{i,K,t}(\tau)=\left[\int_{\min\{\tilde{v}_{i\tau+1}-t,\max\{\tilde{v}_{i\tau}-t,\underline{d}_i\}\}}^{\max\{\tilde{v}_{i\tau}-t,\min\{\tilde{v}_{i\tau+1}-t,\bar{d}_i\}\}} F_i(x)^{\frac{\theta}{1-\theta}}dx\right] \quad (52)$$

and $\tilde{v}_{i\tau}$ is defined as the time gap between the time of the introduction of the τ technology and the efficient vintage at time 0, i.e. $\left(v_{i\tau}-v_{i0}^{*}\right)$, and the integration limits in the ξ terms denote the vintages used that are associated to the technology τ in sector i at time *t*.[6]

Adoption disparities and cross country per capita income differences

Our empirical goals in this paper are threefold. First of all, we aim to compare the theoretical adoption patterns generated by our model with those observed in the data for a broad range of major technologies. Secondly, we want to measure the TFP component that is associated with the adoption of new technologies and understand what determines the observed cross-country variation in this component. Finally, we try to understand what part of the productivity disparities across countries is the result of disparities in the rate at which they adopt new technologies.

To pursue these last two questions, it proves useful to decompose the sector level TFP (Z_{it}) as follows:

$$\ln Z_{it}=\ln a_i+g_i t+g v_{i0}^{*}+\frac{1-\theta}{\theta}\ln\left[\int_{\underline{d}_i}^{\bar{d}_i}F_i(x)^{\frac{\theta}{1-\theta}}dx\right] \quad (53)$$

Throughout, we assume that long run growth in each sector is the same across countries, such that there is no cross-country variation in g_i. Furthermore, we assume that technologies across countries are identical, except for a_i, v^*_{i0}, and the function $F_i(x)$. Hence, for our empirical analysis we impose that θ is also constant across countries.[7]

These two assumptions imply that the percentage difference in sectoral productivity levels between two countries can be approximated by

$$\begin{aligned}\ln Z_{it}-\ln Z_{it}' &= \left(\ln a_i-\ln a_i'\right)+g_i\left(v_{i0}^{*}-v_{i0}^{*\prime}\right)+\frac{1-\theta}{\theta}\left(\ln\left[\int_{\underline{d}_i}^{\bar{d}_i}F_i(x)^{\frac{\theta}{1-\theta}}dx\right]-\ln\left[\int_{\underline{d}_i'}^{\bar{d}_i'}F_i'(x)^{\frac{\theta}{1-\theta}}dx\right]\right)\\ &=\left(\ln a_i-\ln a_i'\right)+\left(\ln Z_i^p-\ln Z_i^{p\prime}\right)\end{aligned} \quad (54)$$

The second part of this term is the part that is attributable to the differences in cross-country technology adoption patterns, i.e. to the differences in the ranges of vintages of technologies that countries employ in production.

[6] See the appendix for the details on the integration limits. Here $v_{i\tau}=-\infty$ if the technology does not have a predecessor and $v_{i\tau+1}=\infty$ if it does not have a successor (yet).

7 Equation (54) shows that our model implies that there are persistent cross-country disparities in sectoral productivity levels. This implication is consistent with some of the evidence presented in Bernard and Jones (1996a, 1996b). They suggest that there is little evidence of convergence in industry productivity levels, especially for manufacturing.

For the interpretation of the results to follow, it is actually worthwhile to consider the interpretation of the above decomposition. Figure 4 illustrates this decomposition graphically. The model implies a distribution of capital for technology *i*, K_{ivt}, over a time varying support of technology vintages, given by $\left[\underline{v}_{it}, \overline{v}_{it}\right]$. This distribution is depicted in the figure. There are three aspects of this distribution that are directly reflected in the TFP measure $ln(Z_{it})$.

These aspects are indicated by (*I*), (*II*), and (*III*) in the figure. They correspond to the following decomposition

$$\ln Z_{it} = \underbrace{\ln a_i}_{(I)} + \underbrace{\frac{1-\theta}{\theta}\ln\left[\int_{\underline{d}_i}^{\overline{d}_i}\left(\frac{F_i(x)}{\exp(g_i \underline{d}_i)}\right)^{\frac{\theta}{1-\theta}} dx\right]}_{(II)} + \underbrace{g_i \underline{v}_{it}}_{(III)}, \text{ where } \underline{v}_{it} = v_{i0}^* + \underline{d}_i \tag{55}$$

Part (*I*) of this decomposition reflects specific TFP differences that are unrelated to the technology adoption decision that is central to this paper. It essentially determines the height of the distribution and can thus be interpreted as the determinant of capital deepening in this model. Parts (*II*) and (*III*) are related to the range of technology vintages in use at each time. Part (*II*) reflects the width of the range of technologies in use, because of this we refer to it as the capital widening part. Because of the CES structure of the production function in our model, the marginal product of each capital vintage is infinite at zero. Therefore TFP is increasing in the degree of diversification of production among different technology vintages. Finally, part (*III*) is determined by the location of the range of vintages, which we identify by its lowerbound. Because this part is identified by the least productive vintage used, we refer to it as capital dragging. That is, the less advanced the vintages that are dragged along, the lower total factor productivity.

The capital deepening part, i.e. (*I*), on the one hand, corresponds closely to the traditional concept of TFP in the neoclassical growth model. It contains all factors that which affect the productivity of all capital in place in the same way. The capital widening and dragging parts, i.e. (*II*) and *(III)*, on the other hand, are specifically identified by the technology adoption structure with which we extended the model in this paper. It is exactly the magnitude of the disparities in capital widening and dragging across countries that is our focus.

All the parameters that are necessary to calculate parts (*II*) and (*III*) of (55) can be estimated by fitting the observed adoption patterns. Hence, our estimation results serve three purposes. First of all, they enable us to consider how well our model fits the observed technology adoption patterns. Secondly, they allow us to infer the sectoral productivity disparities due to technology adoption differences and to understand the role of capital widening and capital dragging in this TFP component. Finally, they permit us to compare the implied measure of TFP disparities due to technology adoption differences with the sector level cross-country productivity differences.

6. Estimation

We estimate the parameters of our model separately for each of major technology categories / country combination using the Generalized Method of Moments (GMM). Our estimation procedure is an application of GMM with nonlinear deterministically trending variables, considered in Andrews and McDermott (1995). In this section we explain how we apply this procedure to our model. The resulting method is a hybrid of reduced form and structural estimation. After calibrating a unique set of values for g, θ, and ζ across time, countries, and technology groups, we only structurally estimate the parameters that are relevant for accounting for the productivity differences due to adoption disparities. The other parameters are estimated in reduced form.

Functional form of F(.)

Since our estimation procedure is parametric, it requires us to make a specific assumption about the functional form of the function *F(.)*, which determines the vintage productivity profile. In principle, any functional form that satisfies the properties described in (41) can be used.

In practice, however, it is important to choose a functional form that satisfies three main criteria. These are (*i*) all relevant parameters needed for the accounting exercise of equation (54) are separately identified, (*ii*) the functional form is flexible enough to fit the wide variety of adoption patterns in the data, (*iii*) the functional form is parsimonious in its parameterization and all parameters have an economic interpretation.

There are many functional forms for *F(.)* that might fit this bill. The one that we will use for our analysis is closely related to the logistic function and equals

$$F_i(x) = 2\frac{\exp(g_i x)}{1+\exp((1+\psi_i)g_i \max\{0,x\})}\text{, where } 0<\psi_i<1 \tag{56}$$

Here g_i is the growth rate of Z_i, while ψ_i determines the rate at which the productivity profile falls below the feasibility frontier for $x>0$. This can be seen by considering the percentage change of *F(.)* with respect to x for $x>0$. This growth rate equals

$$\frac{\partial \ln F_i(x)}{\partial x} = g_i\left[1-(1+\psi_i)\frac{\exp((1+\psi_i)g_i x)}{1+\exp((1+\psi_i)g_i x)}\right] \tag{57}$$

Model specification

For each technology/country combination we observe a time series that reflects the technology adoption pattern for the technology vintages in the particular country. The model that we estimate for each technology category / country combination, indexed by i,j, thus has a panel data structure. Each cross-sectional observation is made up of a technology vintage, indexed by τ.

Depending on whether we have data on the vintage output to GDP ratio or the vintage capital to output ratio, the observed time series reflect $m_{i,Y,j,t}(\tau)$ or $m_{i,K,j,t}(\tau)$. Equation (51) implies that, in reduced form, we can write

$$\ln m_{i,Y,j,t}(\tau) = \beta_{0,i,Y,\tau,j} + \beta_{1,i,Y,\tau,j}t + \ln \xi_{i,Y,j,t}(\tau) \tag{58}$$

and

$$\ln m_{i,K,j,t}(\tau) = \beta_{0,i,K,\tau,j} + \beta_{1,i,K,\tau,j}t + \beta_{2,i,K,\tau,j}\ln Y_{j,t} + \ln \xi_{i,K,j,t}(\tau) \tag{59}$$

These are the equations that we will estimate for each technology category.

In principle, the reduced form coefficients, i.e. the β's, depend on the structural parameters[8]. We will, however, ignore this dependence and will focus our estimation efforts on obtaining estimates of the parameters that determine the nonlinear deterministic trends $ln\xi_{i,Y,j,t}(\tau)$ and $ln\xi_{i,K,j,t}(\tau)$. These are the same parameters that are needed to account for sectoral productivity disparities, following equation (54).

In particular, for each technology category / country combination, we will estimate the the time gap between the time of the introduction of the first technology (in the category) and the efficient vintage at time 0 ($\widetilde{v}_{i\tau} \equiv v_{i,1} - v^*_{i,j,0}$) and the curvature of the function that determines the efficiency of each vintage ($\psi_{i,j}$). We do so conditional on the calibrated values of the parameters θ, g_i, and ζ_i, as well as the transformed technology cut-off vintages $v_{i,\tau} - v_{i,1}$ for the various technology vintages, $\tau>1$.

The estimates of the parameters $\overline{d}_{i,j}$, and $\underline{d}_{i,j}$ are implied by the restriction that equilibrium entry condition holds. This restriction reads

$$F_i\left(\underline{d}_{i,j}\right) = F_i\left(\overline{d}_{i,j}\right) = \left(\frac{\zeta_i}{1-\theta}\right)\left[\int_{\underline{d}_{ij}}^{\overline{d}_{ij}} F_{ij}(x)^{\frac{\theta}{1-\theta}} dx\right]^{\frac{1-\theta}{\theta}} \tag{60}$$

It defines $\psi_{i,j}$ as an implicit function of θ, g_i, $\overline{d}_{i,j}$, and $\underline{d}_{i,j}$. We denote this function as

$$\overline{d}_{i,j} = \overline{D}(\psi_{i,j}; g_i, \theta) \text{ and } \underline{d}_{i,j} = \underline{D}(\psi_{i,j}; g_i, \theta) \tag{61}$$

All of this together implies that the equations we estimate are of the form (58) and (59), where the nonlinear deterministic parts equal[9]

$$\ln \xi_{i,Y,j,t}(\tau) = \ln\left[\int_{\max\{\widetilde{v}_{i\tau}-t, \underline{D}(\psi_{i,j};g_i,\theta)\}}^{\min\{\widetilde{v}_{i\tau+1}-t, \overline{D}(\psi_{i,j};g_i,\theta)\}} \left[\frac{\exp(g_i x)}{\exp((1+\psi_{i,j})g_i \max\{0,x\})}\right]^{\frac{1}{1-\theta}} dx\right] \tag{62}$$

and

$$\ln \xi_{i,K,j,t}(\tau) = \ln\left[\int_{\max\{\widetilde{v}_{i\tau}-t, \underline{D}(\psi_{i,j};g_i,\theta)\}}^{\min\{\widetilde{v}_{i\tau+1}-t, \overline{D}(\psi_{i,j};g_i,\theta)\}} \left[\frac{\exp(g_i x)}{\exp((1+\psi_{i,j})g_i \max\{0,x\})}\right]^{\frac{\theta}{1-\theta}} dx\right]. \tag{63}$$

---

[8] The reduced form parameters β_0 also depend on, among others, units of measurement. If, besides quality growth, one would also include (constant) price declines of units of output and capital in the model (like models of investment specific technological change) then β_1 would also depend on the relative price declines/increases. In that case, g_i is the combined growth rate of embodied and investment specific technological change. Even though our model only includes embodied technological change, in our application we will be agnostic about the relative importance of embodied versus investment specific technological change. Hence, we will not restrict β_1 based on our model.

[9] The limits of the integrals in these two equations simplify because we only consider the case in which $\xi_{i,Y,j,t}(\tau)$ and $\xi_{i,M,j,t}(\tau)$ are strictly positive. This is, by definition, the case for the observations in our data set.

The estimation of this model requires the joint estimation of all the non-linear equations in the panel under several cross equation restrictions. This is the case because both $\psi_{i,j}$ and $\left(v_{i,1}-v^{*}_{i,j,0}\right)$ are constant for all technology vintages in a technological group.[10]

Moment conditions and estimation

For the estimation of our model we use GMM. Our model is one with a nonlinear deterministic trend in the form of $\xi_{i,Y,j,t}(\tau)$ and $\xi_{i,K,j,t}(\tau)$. The properties of GMM in case of nonlinear deterministically trending variables are derived in Andrews and McDermott (1995). They prove that, in this case, GMM parameter estimates are both consistent as well as asymptotically normally distributed.

The application of GMM requires the choice of a set of moment conditions that are used to identify the parameters. Since our model is fully deterministic, it does not contain any structural stochastic components that can be used to define the conditions. Instead, we define the moment conditions in terms of the deviations of the observed variables from their fitted values, i.e. the residuals of the equations. In this sense, our empirical methodology amounts to advanced trend-fitting. It is very similar in nature to the empirical analysis in Laitner and Stolyarov (2004).

Define the residuals $u_{i,Y,j,t}$ and $u_{i,K,j,t}$, respectively, as

$$u_{i,Y,j,t}(\tau)=\ln m_{i,Y,j,t}(\tau)-\beta_{0,i,Y,\tau,j}-\beta_{1,i,Y,\tau,j}t-\ln \xi_{i,Y,j,t}(\tau) \tag{64}$$

and

$$u_{i,K,j,t}(\tau)=\ln m_{i,K,j,t}(\tau)-\beta_{0,i,K,\tau,j}-\beta_{1,i,K,\tau,j}t-\beta_{2,i,K,\tau,j}\ln Y_{j,t}-\ln \xi_{i,K,j,t}(\tau) \tag{65}$$

then, for each country, we assume the existence of a set of instruments $z_{1,j,t},\ldots,z_{r,j,t}$ for which

$$E\left[u_{i,Y,j,t}z_{s,j,t}\right]=E\left[u_{i,M,j,t}z_{s,j,t}\right]=0 \text{ for } s=1,\ldots,r \tag{66}$$

One underlying assumption is important to note here. That is that our application of GMM is based on assumption that $u_{i,Y,j,t}$ and $u_{i,K,j,t}$ are zero mean stationary processes. Since the logarithm of GDP per capita, i.e. $\ln Y_{j,t}$, is part of the equation for $\ln m_{i,K,j,t}(\tau)$ this means that we implicitly assume that the logarithm of GDP per capita is trend stationary.

Our GMM estimates of the parameters are the values for $\left(v_{i,1}-v^{*}_{i,j,0}\right)$'s, and the curvature parameters $\psi_{i,j}$ that minimize a quadratic objective of the sample approximations of the moment conditions of equation (66). The form of the GMM objective, the structure of the minimization problem, as well as the way in which the results of Andrews and McDermott (1995) are applicable to this model, are explained in detail in Appendix B.

[10] Note that $\left(v_{i,1}-v^{*}_{i,j,0}\right)$ is the actual component estimated because $\tilde{v}_{i,j\tau}$ can be decomposed in the following way $\tilde{v}_{i,j\tau}=\left(v_{i,\tau}-v_{i,1}\right)+\left(v_{i,1}-v^{*}_{i,j,0}\right)$ where $\left(v_{i,\tau}-v_{i,1}\right)$ denotes the number of years after the discretization point of the first vintage the discretization point of the τ vintage falls and remember that we use information about the invention of the different technologies to calibrate these terms.

Identification

Before we present the estimates of the structural parameters ψ_i and $\left(v_{i,1} - v^{*}_{i,j,0}\right)$, it is important to consider which features of the data allow us to identify them. The inclusion of an intercept and trend in (58) and (59) means that what is left for ψ_i and $\left(v_{i,1} - v^{*}_{i,j,0}\right)$ to explain is the common non-linear deterministic trend part in the adoption curves of the technology vintages in each technology group.

What the model and the data have in common is that the non-linearity of the adoption curves has two very distinct features. These two features are best illustrated using an empirical example. For this reason, consider the actual, as well as the fitted, adoption curves for telephones and telegrams in the U.S., which are depicted in Panel B of Figure 5.

The first obvious property of these adoption curves is their curvature, both in the early adoption phase of a technology vintage, like that for telephones, as well as in the phase in which a technology vintage declines, like that for telegrams. This curvature is what determines our estimate of ψ_i. The second feature is where the curvature starts in time. This is what determines the estimate of $\left(v_{i,1} - v^{*}_{i,j,0}\right)$.

It is important to emphasize that both of these structural parameters (ψ_i and $\left(v_{i,1} - v^{*}_{i,j,0}\right)$) are the same for all adoption curves fitted within the same technology group. Hence, the fitted curvatures for the curves for both telegrams, telephones, as well as cellphones, are all based on the same value of ψ_i.

7. Empirical results

This section is the culmination of this paper in the sense that we present estimates of the model introduced obtained using an extensive historical dataset on the adoption of several major technologies across countries, spanning almost two centuries. Our focus is on two particular aspects of the results. First of all, how well does the model fit the adoption patterns for the broad range of technologies for which we have data? Does the model capture the common parts of the adoption patterns observed in the data? Secondly, what does the model imply for the productivity disparities across countries that are caused by them using different ranges of technology vintages?

We present our results in four parts. In the first part we describe the data used for the estimation. In the second part we focus on the first aspect of our results, namely the parameter estimates and the quality of the fit of the model. In the third part we discuss what our results reveal about TFP disparities due to differences in the types of technologies used across countries. Finally, in the fourth part we summarize what we consider the bottom line of our results and some of their limitations.

Data

For our empirical analysis we use data from a unique data set, the Historical Cross-Country Technology Adoption Data set, that we assembled and explained in detail in Comin and Hobijn (2003). The data that we use here contain both aggregate measures of economic growth as well as measures of technology adoption for 17 major technologies. The data cover the period 1820-2000 and a sample of twenty-one of the World's leading industrialized countries. Table 1 lists the countries and technologies covered in the data.

As can be seen from Table 1, we identify seven main technology categories, identified by *i*. These technology categories range from textiles, for which we have data on the adoption of spindles, to telecommunications. Each category contains two or more technology vintages, indexed by τ. The measures for these vintages range from steel output, measured in tons, produced in certain types of steel furnaces, to the capital stock of passenger cars in each country.

Besides the index τ and the description of the vintage, Table 1 contains two more pieces of information for each vintage. It lists the 'type' of the measure for the vintage. In the 'type' column, *K* identifies a capital stock measure and *Y* indicates an output measure. The $(v_{i\tau}-v_{i1})$ column denotes how many years after the discretization point of the first vintage the discretization point of vintage τ falls. As described above, we keep these interval lengths fixed across countries during our estimation. The way we have calibrated them is by setting them equal to the difference between the invention dates of the relevant vintage and the first vintage. We used the invention dates listed in Figure 4 of Comin and Hobijn (2004). For sailships we assume that they make up the whole merchant fleet at the beginning of our sample and we do not have a vintage starting date.

Parameter estimates and fit

We start off the presentation of our parameter estimates with a discussion of the parameters that we actually do not estimate. That is, the results that we present in the following are conditional on a set of calibrated parameters, namely g, θ, and ζ. We hold these three parameters constant for all our estimates, across time, countries, and technology groups.

We calibrate g to equal 1.4% annually, which is the growth rate of TFP consistent with a long-run average growth rate of real GDP per capita of 2% and is also approximately equal to the average post-war growth rate of aggregate TFP for the U.S. economy.

Calibration of the parameters θ and ζ is not as straightforward. Our results here are based on the choice of $\theta=0.9$ and $\zeta=0.5$. The combination of these two parameter choices implies an aggregate profit margin, net of fixed operating costs, equal to 5%. This is roughly consistent with the average operating margin reported for publicly traded firms in the U.S. over the last three decades. A sensitivity analysis of our results suggests that they do not change dramatically for θ between *0.8* and *0.95*. For operating costs higher than 5% of nominal output, i.e. $\zeta>0.5$, we found that the model would predict much shorter technology lifecycles than we observe in the data.

The gist of what the actual data looks like and how closely the model fit the data is contained in Figure 5. It displays the fit and actual adoption curves for steel in the U.S. (panel A.), telephones and telegrams in the U.S. (panel B), and passenger transportation for the U.S. and the U.K. (panels C. and D.). The fit depicted for the four examples in Figure 5 is actually very representative for that of the other country / technology group combinations.

As can be seen from this figure, the model seems to fit quite well the adoption patterns of a broad range of technologies well. Given the 'one size fits all' nature of our estimation in which we apply our one model to a wide variety of technology groups, this consistent quality of fit implies that our structural model manages to successfully replicate the common general historical adoption patterns of major technologies. Further, the fact that we obtain a good fit despite all the restrictions imposed (both in the functional forms and in the calibrated parameter values) means that these restrictions cannot be rejected in a statistical sense.

So what are the parameter estimates that yield such a fit, and do they make economic sense? To answer this question, consider Table 2. This table lists the estimates of ψ_i and $\left(v_{i,1} - v^*_{i,j,0}\right)$ as well as the associated standard errors for all country / technology group combinations.

The parameter estimates confirm the general implications of the model. That is, for technology groups where the adoption of new vintages took relatively long, like the shipping and the other two transportation technologies, we obtain low estimates of ψ_i, generally between 0.3 and 0.6. Technologies with more rapid adoption of vintages, like mass communication and steel, yield much higher estimates of ψ_i, generally between 1 and 1.5. For 'early' technologies, like textiles, we obtain low estimates of $\left(v_{i,1} - v^*_{i,j,0}\right)$, presumably in large part because $v_{i,1}$ in that case is much smaller than for 'later' technologies, like telecommunications and transportation. A notable exception is shipping, where the steamboat was invented in the early part of our sample but our estimates of $\left(v_{i,1} - v^*_{i,j,0}\right)$ are relatively large. Combined with the associated low estimates of ψ_i this suggests that it takes a very long time of shipping technologies to fully mature.

It turns out that our parameter estimates are not particularly sensitive with respect to the technologies included in the technology groups. That is, where the grouping of the technologies for textiles, steel, and shipping is fairly uncontroversial, the grouping of the technologies in the other categories is less so. However, eliminating cell phones from the telecommunications group or dropping air transportation from the two transportation categories does not really influence the result. This is because the parameters are mainly identified by the curvature in the adoption patterns of the other technology vintages.

The standard errors of the estimates are revealing. Andrews' and McDermott's (1995) GMM method is very unforgiving in terms of the implied standard errors. In cases where the model does not yield a good fit to the data, the standard errors are a factor ten higher than in case of a satisfactory fit. As the standard errors show the 'one size fits all' character of our empirical analysis yields results that are best characterized as 'one size fits many'.

For most country / technology combinations we obtain fairly reliable parameter estimates. However, in some cases the data do not allow us to allow for accurate inference on ψ_i and $\left(v_{i,1} - v^*_{i,j,0}\right)$. Further analysis of our data reveals that there are basically four reasons why this might be the case.

The first is simply a lack of data. This is the case for steel in Greece, Ireland, and Portugal, where we do not have data on open-hearth steel. There might be two reasons for this lack of data. On the one hand, they might simply not have been collected. On the other, because of the relatively small size of the steel manufacturing industries in these countries, open hearth steel might actually not have been produced.

The second reason relates to the latter point above. Sometimes data is available but based on a very small sample. This is the case for Danish steel production, which shows several kinks in the adoption curve of steel. These kinks presumably correspond to new furnaces coming online. The indivisibility of investment in such furnaces, combined with the small size of the Danish steel producing sector, thus leads to a failure of our model to fit the observed steel adoption pattern.

The third reason is that some of our data is so noisy that the noise dominates the second order curvature that identifies our structural parameters. This seems to be the case, for example, for the communications data for Greece.

Finally, the annual frequency of observation also implies that we have a hard time estimating the curvature parameter ψ_i in case of a relatively rapid adoption of a technology. If a technology is adopted rapidly then the curvature only spans very few annual observations. Consequently, very few observations contribute to the identification of ψ_i. It turns out that in those cases the estimation method tends to ignore these observations and veers towards fitting a flat adoption curve where ψ_i is small. This seems to be the case for tele- and mass communications in Canada.

However, in spite of the limited number of cases where we do not obtain reliable estimates of the structural parameters, the gist of our results seems to be that for the vast majority of country / technology group combinations in our data set we are able to obtain fairly reliable estimates of ψ_i and $\left(v_{i,1} - v^*_{i,j,0}\right)$.

Productivity disparities

The estimates of the structural parameters also allow us to calculate estimates of the part of TFP determined by the adoption of technologies, i.e. $ln(Z^p_{ij})$, as well as the range of technologies used, i.e. $\left[\underline{v}_{ijt}, \bar{v}_{ijt}\right]$.[11] Table 3 lists the implied values of $\left[\underline{v}_{ijt}, \bar{v}_{ijt}\right]$ for each of the country / technology group combinations. Table 4 contains the associated levels of $ln(Z^p_{ij})$ and their standard errors.

In order for our analysis of the productivity estimates to not be influenced by large estimation errors, we have chosen to focus on the ones that are relatively reliably estimated. The criterion we use is a standard error of $ln(Z^p_{ij})$ of 0.1 or less[12]. The parameter estimates for which this is the case are shaded gray in Table 4.

The columns labeled $ln(Z^p_{ij})$ in Table 5 contain the two main descriptive statistics for our productivity measure, namely the number of countries for which we obtain a reliable estimate, i.e. *N*, and the standard deviation of $ln(Z^p_{ij})$ across these countries. The standard deviation of $ln(Z^p_{ij})$ turns out to be between 9%, for the limited number of countries for which we have estimates for shipping, and 25%, for the 13 countries for which we have estimates for mass communications. The 1990 standard deviation of the levels of log real GDP per capita for the corresponding countries varies from 9% to 19%, depending on the technology group. In this sense, we find that technology adoption induced disparities in TFP are of a magnitude similar to the disparities in real GDP per capita.

A more careful investigation of the rankings of the technology adoption component of TFP yields a second interesting observation: there is no uniformity in the country that leads in this component across technologies. The US leads in textiles and mass communications. France leads in shipping and steel where has the same adoption component of TFP as Italy. In telecommunications both Australia and Portugal are ahead in the adoption component of TFP, while Italy and Belgium are the leaders in cargo and passenger transportation, respectively.

---

[11] The estimates of $ln(Z^p_{ij})$ are very robust to the calibration of *g*.

[12] Changing this cut-off point to 0.2 does change some of the details of the results here but not the nature of their implications.

After this descriptive analysis of $ln(Z^p_{ij})$, one may wonder what forces drive cross-country disparities in the adoption component of TFP. To address this question, consider the following decomposition of $ln(Z^p_{ij})$ into capital widening and dragging based on (55):

$$\ln Z_{ij}^{p} = \underbrace{\frac{1-\theta}{\theta}\ln\left[\int_{\underline{d}_{ij}}^{\bar{d}_{ij}}\left(\frac{F_i(x)}{\exp\left(g_i \underline{d}_{ij}\right)}\right)^{\frac{\theta}{1-\theta}}dx\right]}_{\text{widening}} + \underbrace{g_i\left(v_{i1}-v_{ij0}^{*}\right)}_{\text{dragging}} \tag{67}$$

This decomposition is closely related to the technology usage intervals $\left[\underline{v}_{ijt}, \bar{v}_{ijt}\right]$, listed in Table 3.

What we find, for all technologies in our dataset, is that there is a virtually perfect correlation between the lower bound of the vintages used in production, $\underline{v}_{ijt}$, and the adoption component of TFP ($ln(Z^p_{ij})$). Since differences between $\underline{v}_{ijt}$ across countries reflect differences in capital dragging, what seems to matter most for $ln(Z^p_{ij})$ is not when a country starts using a technology but when it actually stops using it. Why is this the case in our model? The are two main reasons that it is mostly capital dragging that determines $ln(Z^p_{ij})$ and not capital widening.

First of all, the widths of the technology ranges in use are not that different across countries. To see this, consider Table 3. This table shows that there tend to be fairly large variations in $\underline{v}_{ijt}$ and $\bar{v}_{ijt}$, both within and between technology groups. In spite of the large differences in $\underline{v}_{ijt}$ and $\bar{v}_{ijt}$ across countries, the variation in the widths of $\left[\underline{v}_{ijt}, \bar{v}_{ijt}\right]$ across countries is a lot less. For those countries for which we have reliable estimates, the average duration of use of a technology vintage is 81 years for textiles, 71 years for steel, 155 years for shipping, 120 years for telecommunications, 61 years for mass communication, and 120 and 126 years for cargo and passenger transportation[13].

Secondly, because of our calibration of $\theta=0.9$, the CES aggregate of the capital widening part does not depend a lot on the tails. Instead, it mainly depends on the highest values of the function being aggregated. These values turn out to be almost constant across countries and thus so is the capital widening part. Lower values of θ would yield more variations in the capital widening parts. However, they would also imply unrealistically high markups and distortions.

Finally, the last important question that remains to be addressed is about the relevance of the adoption component of TFP estimated for understanding sectoral differences in TFP across countries. To answer this question, we compare our estimates of the sectoral adoption component of TFP ($ln(Z^p_{ij})$) with the 1985 measures

[13] It is important here to emphasize that these are technology vintage lifecycle lengths. For example, the 61 years for mass communication implies that 61 years after the first type of black and white television was first used, no one will use that type of black and white TV anymore. This is a very different concept from the average age of these black and white televisions. This is because the vintage specific capital stock, K_{ivt}, changes over time.

[14] Correlations between our estimates of $ln(Z^p_{ij})$ and the estimates of aggregate 1985 TFP levels calculated in Klenow and Rodríquez-Clare (1997) yield results very similar to those presented for real GDP per capita here. This is because the aggregate TFP levels and real GDP per capita are highly correlated.

of sectoral TFP computed by Bernard and Jones (1996a) from the International Sectoral Data Base (ISDB). We will refer to these measures as $ln(Z_{ij1985})$.

The results of this comparison are in Table 5. Even though the sectors in the data set used by Bernard and Jones (1996a) do not match up perfectly with ours and even though some of the capital goods that we include in our analysis, like passenger cars, are used mostly for home production, we find some remarkably high correlations between the estimates of $ln(Z^p_{ij})$ and $ln(Z_{ij1985})$. For shipping, cargo and passenger transportation technologies we find correlations of 0.65 and higher. In spite of the fact that the spindle technology only constitutes a small fraction of the production technology for textiles, we find a 0.36 correlation between its $ln(Z^p_{ij})$ and $ln(Z_{ij1985})$ for the textiles sector. Surprisingly, we find a small negative correlation between $ln(Z^p_{ij})$ and $ln(Z_{ij1985})$ for steel. For the communications sectors we do not find much of a correlation. However, for those sectors we measure the adoption of capital vintages by users rather than by the providers of the services. Hence, in those cases Bernard and Jones' (1996a) sectoral TFP levels might not be such a great proxy for $ln(Z_{ijt})$ in our model.

What the $ln(Z^p_{ij})$ and $ln(Z_{ij1985})$ have in common across technologies is that the cross-country variation in the ISBD measures of TFP ($ln(Z_{ij1985})$) for all technologies are substantially higher than for adoption component of TFP estimated in this paper. The ratio in the cross-country standard deviations ranges from 1.25 for shipping to 11 for telecommunications. This difference in the variance between the adoption component and the ISBD measures of TFP is probably due to the combination of two factors. First, the ISBD data is quite noisy. One of the sources of this noise is that, due to data limitations, the same PPP adjustment is applied to the capital stock in all sectors, irrespective of the capital composition. As a result, the estimate of the variance of sectoral TFP that results is biased upwards and, presumably, the correlation between $ln(Z^p_{ij})$ and $ln(Z_{ij1985})$ that we report in Table 5 is biased downward. Secondly, at a sectoral level there seems to be a positive correlation between capital deepening on the one hand and capital widening and dragging on the other.

Bottom line

After diving into the detailed results, there are three main conclusions that we can take away from them. First and foremost, our theoretical framework seems to be able to replicate the general common characteristics of the adoption patterns of several major technologies over the past two centuries. Secondly, what matters most for the influence of technology adoption on TFP is actually not when a technology is first used but when it is last used. That is, how many old technology vintages are being dragged along. Finally, the technology adoption component of TFP tends to be positively correlated to sectoral TFP levels in the cross-country dimension, though its variance is substantially smaller.

It is important to bear in mind, though, that these results are obtained using structural estimation and historical data for a limited set of advanced industrialized economies.

The application of structural estimation here allows us to empirically identify the capital widening and dragging concepts that provide us with a new insight into how technology adoption affects TFP. The structure that we imposed, however, in large part guides our results. Further analysis of what restrictions on the structure

[15] It is important here to emphasize that these are technology vintage lifecycle lengths. For example, the 61 years for mass communication implies that 61 years after the first type of black and white television was first used no one will use that type of black and white TV anymore. This is a very different concept from the average age of these black and white televisions. That is, the capital stock, $K_{i\tau t}$, changes over time.

can be relaxed without affecting the identification of capital widening and dragging will allow us to verify the theoretical robustness of our results.

By limiting our analysis to the set of, relatively similar, leading industrialized economies, we have biased our results to finding little differences in adoption patterns and disparities across countries. In this sense, our analysis here is subject to the same criticism that DeLong (1988) had against Baumol (1986). An extension of our analysis, to more countries and technologies, requires the use of data that has not been collected yet. When they become available, combining such data with the model and estimation method introduced here would allow us get more insight into the technological gap between developing and industrialized economies. Some of the theories on technology adoption disparities, like Basu and Weil's (1998) theory of appropriate technologies, seem to be especially aimed at explaining this supposed gap. In this respect, it is interesting to note that, at the beginning of our sample, some of the, ex-post, industrialized countries in our sample were actually in a similar stage of economic development as many developing economies now. However, even for the technologies adopted in the early part of our sample, like railroads and textiles, we do not find large adoption disparities.

8. Conclusion

In this paper, we introduced a growth model where different firms simultaneously produce intermediate goods with varying productivity levels. New intermediate goods are potentially more productive but are, initially operated below their full-efficiency. The array of intermediate goods adopted for production is endogenous. The combination of these simple elements yields large returns: At the micro level, our model is a model of slow diffusion of technologies (embobied in new intermediate goods) qualitatively consistent with the stylized facts: S-shaped diffusion curves and investment in old vintages. The model responds to Prescott (1997)'s demand: "Needed: A Theory of TFP". The endogenous adoption of new technologies and the variety externalities provide a theory of TFP both at the sector and aggregate level. Further, our model does all this within a well-known context because at the aggregate level our model is observationally equivalent to the Neoclassical growth model.

By combining this model with the Historical Cross-Country Data set introduced in Comin and Hobijn (2004) we are able to bridge an endemic disconnect between theory and data in the technology diffusion literature. The estimation of the structural parameters that determine the adoption patters, yields three main insights. Most importantly, cross-country adoption patterns of major technologies over the past two centuries exhibit very general common characteristics that can be explained very well by a model that at the aggregate mimics the standard neoclassical growth model. Secondly, the effect of technology adoption on TFP seems to be mainly determined by a concept that we call capital dragging. Countries that still use older technologies seem to have lower productivity levels attributable to technology adoption. On the other hand, the width of the range of technologies in use and the most modern technology in use seem to be a lot less important determinants of productivity disparities. Finally, cross-country TFP disparities due to differences in technology adoption patterns are positively correlated with overall sectoral TFP levels. However, they seem to be substantially smaller. If this large variance in the TFP measures of Bernard and Jones (1996a) is real, this suggests that the majority of sectoral TFP disparities is due to factors, like regulations and institutions, that affect all technology vintages in the same way.

We leave for future work understanding the channels by which regulations and institutions affect TFP. The determinants of the estimated shape of the efficiency function (ψ_i) will also be studied in subsequent work. Finally, the scope of the theoretical and empirical contributions of this paper far exceed the set of technologies and countries that they are applied to in this paper. They can potentially provide insights into technology adoption patterns and disparities for many more technologies as well as a broader set of countries and regions.

References

Anderson, Simon P., Andre de Palma, and Jacques-Francois Thisse (1992), Discrete Choice Theory of Product Differentiation, Cambridge, MA: MIT Press.

Andrews, Donald W.K., and C. John McDermott (1995), "Nonlinear Econometric Models with Deterministically Trending Variables", *Review of Economic Studies*, 62, 343-360.

Bahk, Byong-Hyong, and Michael Gort (1993), "Decomposing Learning by Doing in New Plants", *Journal of Political Economy*, 101, 561-583.

Basu, Susanto and David N. Weil (1998), "Appropriate Technology and Growth", *Quarterly Journal of Economics*, 113, 1025-1054.

Baumol, William J., (1986), "Productivity Growth, Convergence, and Welfare: What the Long-Run Data Show", *American Economic Review*, 76, 1072-1085.

Bernard, Andrew B., and Charles I. Jones (1996a), "Comparing Apples to Oranges: Productivity Convergence and Measurement Across Industries and Countries", *American Economic Review*, 86, 1216-1238.

Bernard, Andrew B., and Charles I. Jones (1996b), "Productivity Across Industries and Countries: Time Series Theory and Evidence", *Review of Economics and Statistics*, 78, 135-146.

Caselli, Francesco, and Jaume Ventura (2000), "A Representative Consumer Theory of Distribution", *American Economic Review*, 90, 909-926.

Caselli, Francesco, and Wilbur John Coleman (2001), "Cross-Country Technology Diffusion: The Case of Computers", *American Economic Review – AEA Papers and Proceedings*, 91, 328-335.

Chari, V.V., and Hugo Hopenhayn (1991), "Vintage Human Capital, Growth, and the Diffusion of New Technology", *The Journal of Political Economy*, Vol. 99, No. 6. (Dec.), pp. 1142-1165.

Comin, Diego, and Bart Hobijn (2003), "The Historical Cross-Country Technology Adoption Dataset", *mimeo*.

Comin, Diego, and Bart Hobijn (2004), "Cross-country technology adoption: making the theories face the facts", *Journal of Monetary Economics*, Volume 51, 38-83.

De Long, J. Bradford (1988), "Productivity Growth, Convergence, and Welfare: Comment", *American Economic Review*, 78, 1138-1154.

Fernandes, Ana (2003), "What Does the Walrasian Auctioneer Know? Technology Adoption and Financial Innovation", *mimeo*, CEMFI/UPF.

Gort, Michael, and Steven Klepper (1982), "Time Paths in the Diffusion of Product Innovations", *Economic Journal*, 92, 630-653.

Griliches, Zvi (1957), "Hybrid Corn: An Exploration in the Economics of Technological Change", *Econometrica*, 25, 501-522.

Greenwood, Jeremy, Zvi Hercowitz, and Per Krusell (1997), "Long-Run Implications of Investment-Specific Technological Change", *American Economic Review*, 87, 342-362.

Hall, Rober E. Hall, and Charles I. Jones (1999), "Why Do Some Countries Produce So Much More Output per Worker than Others?", *Quarterly Journal of Economics*, Vol. 114, pp. 83-116.

Holmes, Thomas J., and James A. Schmitz Jr. (2001), "A Gain from Trade: From Unproductive to Productive Entrepreneurship", *Journal of Monetary Economics*, 47, 417-446.

Irwin, Douglas A., and Peter J. Klenow (1994), "Learning-by-Doing Spillovers in the Semiconductor Industry", *Journal of Political Economy*, 102, 1200-1227.

Jerzmanowski, Michael (2002), "TFP difference: Appropriate Technology vs. Efficiency", *mimeo*, Brown University.

Johansen, Leif (1959), "Substitution Versus Fixed Production Coefficients in the Theory of Economic Growth: A Synthesis", *Econometrica*, 27, 157-176.

Jovanovic, Boyan, and Yaw Nyarko (1995), "A Bayesian Learning Model Fitted to a Variety of Empirical Learning Curves", *Brookings Papers on Economic Activity*: *Microeconomics*, 247-305.

Klenow, Peter J. and Andrés Rodríquez-Clare (1997), "The Neoclassical Revival in Growth Economics: Has It Gone Too Far?", *NBER Macroeconomics Annual*, B. Bernanke and J. Rotemberg eds., Cambridge, MA: MIT Press, 73-102.

Laitner, John, and Dmitriy Stolyarov (2004), "Technological Change and the Stock Market", *American Economic Review*, 94, 1240-1267.

Levhari, David (1968), "A Note on Houthakker's Aggregate Production Function in a Multiform Industry", *Econometrica*, Vol. 36,No. 1 (Jan.), 151-54.

Lucas, Robert E., Jr. (1990), "Why Doesn't Capital Flow from Rich to Poor Countries?", *American Economic Review*, 80 (Papers and Proceedings), 92-96.

Mansfield, Edwin (1061), "Technical Change and the Rate of Imitation", *Econometrica,* Vol. 29, No. 4. (Oct.), pp. 741-766.

Parente, Stephen L. and Edward C. Prescott (1994), "Barriers to Technology: Adoption and Development", *Journal of Political Economy*, 102, 298-321.

Prescott, Edward C. (1997), "Needed: A Theory of Total Factor Productivity" Federal Reserve Bank of Minneapolis, Staff Report 242.

Saxonhouse, Gary and Gavin Wright (2000), "Technological Evolution in Cotton Spinning, 1878-1933", *mimeo*.

Solow, Robert (1960), "Investment and Technical Progress", in Arrow, Karlin and Suppes (eds.), *Mathematical Methods in the Social Sciences*, Stanford University Press.

A. Mathematical details

Derivation of equations (20) through (26):

Each individual intermediate goods supplier equates the marginal revenue of labor to the real wage rate. For the particular production function and demand function that we use, this boils down to the optimality condition

$$W_t L_{ivt} = (1-\alpha)\theta P_{it} Y_{it}^{1-\theta} Y_{ivt}^{\theta} \tag{69}$$

Similarly, each firm also equates the marginal revenue of capital to its user cost. This means that

$$(r_t + \delta) K_{ivt} = \alpha\theta P_{it} Y_{it}^{1-\theta} Y_{ivt}^{\theta} \tag{70}$$

When we integrate both of these optimality conditions over the firms in the market, we obtain

$$W_t L_{it} = W_t \int_{\underline{v}_{it}}^{\bar{v}_{it}} L_{ivt} dv = (1-\alpha)\theta P_{it} Y_{it}^{1-\theta} \int_{\underline{v}_{it}}^{\bar{v}_{it}} Y_{ivt}^{\theta} dv = (1-\alpha)\theta P_{it} Y_{it}^{1-\theta} \int_{\underline{v}_{it}}^{\bar{v}_{it}} Y_{ivt}^{\theta} dv = (1-\alpha)\theta P_{it} Y_{it} \tag{71}$$

as well as

$$(r_t + \delta) K_{it} = \alpha\theta P_{it} Y_{it} \tag{72}$$

The above two equations are the sectoral optimality conditions introduced in (26).

By dividing the optimality conditions for the individual intermediate goods suppliers by those of the whole sector, we obtain that

$$\frac{K_{ivt}}{K_{it}} = \frac{L_{ivt}}{L_{it}} = \left(\frac{Y_{ivt}}{Y_{it}}\right)^{\theta} \tag{73}$$

Substituting this into the production function of the intermediate goods suppliers, we obtain that

$$Y_{ivt} = Z_{ivt} \left(\frac{Y_{ivt}}{Y_{it}}\right)^{\theta} K_{it}^{\alpha} L_{it}^{1-\alpha} = \left[Z_{ivt} \frac{K_{it}^{\alpha} L_{it}^{1-\alpha}}{Y_{it}^{\theta}} \right]^{\frac{1}{1-\theta}} \tag{74}$$

Integrating this expression over all suppliers in the market yields that

$$Y_{it}^{\theta} = \int_{\underline{v}_{it}}^{\bar{v}_{it}} Y_{ivt}^{\theta} dv = \left[\frac{K_{it}^{\alpha} L_{it}^{1-\alpha}}{Y_{it}^{\theta}} \right]^{\frac{\theta}{1-\theta}} \left[\int_{\underline{v}_{it}}^{\bar{v}_{it}} Z_{ivt}^{\frac{\theta}{1-\theta}} dv \right] \tag{75}$$

Solving for Y_{it}, this equation simplifies to the sectoral production function (20).

Derivation of equations (27) through (30):

The derivation of equations (27) through (30) is mathematically identical to that of (20) through (26). So, we will only show how to derive the initial optimality conditions that are similar to (69) and (70). This can be derived by substituting the final goods demand functions

$$P_{it}Y_{it} = Y_t^{1-\rho}Y_{it}^{\rho} \tag{76}$$

into equation (26). This yields that the factor demands for each sector satisfy

$$W_tL_{it} = (1-\alpha)\theta Y_t^{1-\rho}Y_{it}^{\rho} \text{ and } (r_t+\delta)K_{it} = \alpha\theta Y_t^{1-\rho}Y_{it}^{\rho} \tag{77}$$

which are equivalent to (69) and (70).

The rest of the equations and results follow in the same way as above.

Derivation of equations (36) and (37):

We will derive the second parts of equations (36) and (37) here. The first parts can be derived in a very similar manner. Their derivation just involves the same results, but then one level up the nested CES structure of our model.

Combining the production function, i.e. equation (10), and result (73) we obtain that

$$Y_{ivt} = Z_{ivt}K_{ivt}^{\alpha}L_{ivt}^{1-\alpha} = Z_{ivt}K_{it}^{\alpha}L_{it}^{1-\alpha}\left(Y_{ivt}/Y_{it}\right)^{\theta} \tag{78}$$

Dividing both sides of this equation by Y_{it}, we obtain that

$$Y_{ivt} = \left(Z_{ivt}/Z_{it}\right)\left(Y_{ivt}/Y_{it}\right)^{\theta} \tag{79}$$

which can be rewritten as the second part of equation (36). That is

$$\left(Y_{ivt}/Y_{it}\right) = \left(Z_{ivt}/Z_{it}\right)^{\frac{1}{1-\theta}} \tag{80}$$

The second part of equation (37) follows from applying the above result to the demand function, i.e. equation (8). Doing so yields

$$\left(\frac{P_{ivt}}{P_{it}}\right) = \left(\frac{Y_{it}}{Y_{ivt}}\right)^{1-\theta} = \left(\frac{Z_{it}}{Z_{ivt}}\right) \tag{81}$$

Derivation of equations (51) and (52):

We will limit ourselves to showing the derivation of $m_{i,K,t}(\tau)$, because that one turns out to be the most difficult of the two. The adoption measure $m_{i,K,t}(\tau)$ is defined as

$$m_{i,K,t}(\tau) = \int_{V_i(\tau)} \frac{K_{ivt}}{Y_t} dv \tag{82}$$

From the user cost equation, (70), and the user cost aggregation result, (30), we obtain that

$$\frac{K_{ivt}}{Y_t} = \frac{\alpha\theta}{(r_t+\delta)} P_{ivt}\left(\frac{Y_{ivt}}{Y_t}\right) = \left(\frac{K_t}{Y_t}\right) P_{ivt}\left(\frac{Y_{ivt}}{Y_t}\right) \tag{83}$$

Because the equilibrium labor supply equals one, i.e. $L_t=1$ for all t, we can manipulate the aggregate Cobb-Douglas production function, i.e. (27), to obtain the aggregate capital output ratio in equilibrium

$$\frac{K_t}{Y_t} = \left(Y_t \Big/ Z_t^{\frac{1}{1-\alpha}}\right)^{\frac{1-\alpha}{\alpha}} \tag{84}$$

Combining the above two results with the normalization that $P_t=1$ for all t, gives

$$\frac{K_{ivt}}{Y_t} = \left(\frac{Y_t}{Z_t^{\frac{1}{1-\alpha}}}\right)^{\frac{1-\alpha}{\alpha}} \left(\frac{P_{ivt}}{P_{it}}\right)\left(\frac{P_{it}}{P_t}\right)\left(\frac{Y_{ivt}}{Y_{it}}\right)\left(\frac{Y_{it}}{Y_t}\right) \tag{85}$$

Applying equations (20) through (26) as well as (39) allows us to write this as

$$\frac{K_{ivt}}{Y_t} = \left(\frac{Y_t}{Z_t^{\frac{1}{1-\alpha}}}\right)^{\frac{1-\alpha}{\alpha}} \left(\frac{Z_{it}}{Z_t}\right)^{\frac{\rho}{1-\rho}} \left(\frac{Z_{ivt}}{Z_{it}}\right)^{\frac{\theta}{1-\theta}} = \left(\frac{Y_t}{Z_t^{\frac{1}{1-\alpha}}}\right)^{\frac{1-\alpha}{\alpha}} \left(\frac{Z_{it}}{Z_t}\right)^{\frac{\rho}{1-\rho}} \left(\frac{Z_{iv_{it}^*t}}{Z_{it}}\right)^{\frac{\theta}{1-\theta}} F\left(v - v_{it}^*\right)^{\frac{\theta}{1-\theta}} \tag{86}$$

At time t, $K_{ivt}>0$ only for $v \in V_{it}^{use}$, where

$$V_{it}^{use} = \left\{ v \,\middle|\, v_{it}^* + \underline{d}_i < v < v_{it}^* + \overline{d}_i \right\} \tag{87}$$

Hence, for the calculation of $m_{i,K,t}(\tau)$, we only have to integrate over $V_i(\tau) \cap V_{it}^{use}$. This intersection is given by the condition

$$\left(v - v_{it}^*\right) \in \left(\min\left\{\widetilde{v}_{i\tau+1} - t, \max\left\{\widetilde{v}_{i\tau} - t, \underline{d}_i\right\}\right\}, \max\left\{\widetilde{v}_{i\tau} - t, \min\left\{\widetilde{v}_{i\tau+1} - t, \overline{d}_i\right\}\right\}\right) \tag{88}$$

which corresponds to the set over which we integrate in equation (52).

B. Econometric details

In this appendix we derive the GMM objective function and illustrate how we minimize it with respect to the parameters of the model. We also show what the results of Andrews and McDermott (1995) imply for the statistical properties of our estimates.

In this appendix, we will limit ourselves to the description of the estimation of the model for the case in which the technology adoption measures used are the vintage capital to output ratios. The estimation of the model in case the measures are output ratios or a mix of both output ratios and capital-output ratios involves the same method as described here and is just a simple generalization.

Before we describe the details of the estimation, we first introduce a bit of notation. We will denote the vector with structural parameters that are estimated as

$$\boldsymbol{\theta}_{ij} = \left[\left(\nu_{i1} - \nu_{i0}^{*} \right) \quad \psi_{ij} \right]' \tag{89}$$

It turns out that the estimation is most easily explained in terms of matrix notation. For this reason, we introduce the following matrices

$$\mathbf{m}_{i,\tau,j} = \left[\ln m_{i,K,\tau,j,1} \quad \ldots \quad \ln m_{i,K,\tau,j,T} \right]', \; \boldsymbol{\xi}_{i,\tau,j} = \left[\ln \xi_{i,K,j,1}(\tau) \quad \ldots \quad \ln \xi_{i,K,j,T}(\tau) \right]' \tag{90}$$

$$\mathbf{X}_{i,\tau,j} = \begin{bmatrix} 1 & \ldots & 1 \\ 1 & \ldots & T \\ \ln Y_{i,1} & \ldots & \ln Y_{i,T} \end{bmatrix}', \; \mathbf{Z}_{i,\tau,j} = \begin{bmatrix} z_{1,j,1} & \cdots & z_{1,j,T} \\ \vdots & \ddots & \vdots \\ z_{r,j,1} & \cdots & z_{r,j,T} \end{bmatrix}' \tag{91}$$

and

$$\boldsymbol{\beta}_{i,\tau,j} = \left[\beta_{0,i,K,\tau,j} \quad \beta_{1,i,K,\tau,j} \quad \beta_{2,i,K,\tau,j} \right] \tag{92}$$

Using this notation, we can write equation (59) for all time series observations, i.e. $t=1,\ldots,T$, for the cross-sectional observation τ,j as

$$\mathbf{m}_{i,\tau,j} = \mathbf{X}_{i,\tau,j}\boldsymbol{\beta}_{i,\tau,j} + \boldsymbol{\xi}_{i,\tau,j} \tag{93}$$

The finite sample approximation of the r moment conditions, (66), that forms the identifying backbone of our analysis can be written as

$$\boldsymbol{\mu}_{i,\tau,j} = \frac{1}{T}\mathbf{Z}'_{i,\tau,j}\left(\mathbf{m}_{i,\tau,j} - \mathbf{X}_{i,\tau,j}\boldsymbol{\beta}_{i,\tau,j} - \boldsymbol{\xi}_{i,\tau,j} \right) = \mathbf{0} \tag{94}$$

The problem is that we need to evaluate these sample moment conditions for all cross-sectional observations, i.e. all $\tau=1,\ldots,q$ for which there is data, jointly. For this purpose, it is easiest to stack the individual equations (93) in a way that essentially yields a Seemingly Unrelated Regressions representation of the model.

Doing so requires the definition of yet another set of matrices. These are

$$\mathbf{m}_{i,j} = \left[\mathbf{m}'_{i,1,j} \quad \dots \quad \mathbf{m}'_{i,q,j}\right]', \ \boldsymbol{\xi}_{i,j} = \left[\boldsymbol{\xi}'_{i,1,j} \quad \dots \quad \boldsymbol{\xi}'_{i,q,j}\right]', \ \boldsymbol{\beta}_{i,j} = \left[\boldsymbol{\beta}'_{i,1,j} \quad \dots \quad \boldsymbol{\beta}'_{i,q,j}\right]' \tag{95}$$

$$\mathbf{X}_{i,j} = \begin{bmatrix} \mathbf{X}_{i,1,j} & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{X}_{i,q,j} \end{bmatrix}, \text{ and } \mathbf{Z}_i = \begin{bmatrix} \mathbf{Z}_{i,1,j} & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{Z}_{i,q,j} \end{bmatrix} \tag{96}$$

These matrices allow us to jointly write all the moment conditions associated with our panel data model as

$$\boldsymbol{\mu}_{i,j} = \frac{1}{T}\mathbf{Z}'_{i,j}\left(\left(\mathbf{m}_{i,j} - \boldsymbol{\xi}_{i,j}\right) - \mathbf{X}_{i,j}\boldsymbol{\beta}_{i,j}\right) = \mathbf{0} \tag{97}$$

Estimation of the model requires both the estimation of the reduced form parameters that we have captured in the vector $\boldsymbol{\beta}_{i,j}$ as well as the structural parameters in the vector $\boldsymbol{\theta}_{i,j}$. The estimates of both of these parameter vectors are the values $\hat{\boldsymbol{\beta}}_{i,j}$ and $\hat{\boldsymbol{\theta}}_{i,j}$ that minimize the GMM objective function, $Q_{i,j}$, which is defined as

$$Q_{i,j} = \frac{1}{T^2}\left(\left(\mathbf{m}_{i,j} - \boldsymbol{\xi}_{i,j}\right) - \mathbf{X}_{i,j}\boldsymbol{\beta}_{i,j}\right)'\mathbf{Z}_{i,j}\mathbf{W}_{i,j}\mathbf{Z}'_{i,j}\left(\left(\mathbf{m}_{i,j} - \boldsymbol{\xi}_{i,j}\right) - \mathbf{X}_{i,j}\boldsymbol{\beta}_{i,j}\right) \tag{98}$$

where the matrix $\mathbf{W}_{i,j}$ is the moment weighting matrix. The weighting matrix that we use is

$$\mathbf{W}_{i,j} = \left(\mathbf{Z}'_{i,j}\mathbf{Z}_{i,j}\right)^{-1} \tag{99}$$

which would be optimal if the errors to the equation are uncorrelated and homoskedastic. In practice, however, the weighting matrix does not matter much for our estimation results. Results using an identity weighting matrix are very similar to the ones presented in this paper.

To facilitate the numerical minimization of $Q_{i,j}$ with respect to the elements of $\boldsymbol{\theta}_i$, it is convenient to concentrate the reduced form parameters $\boldsymbol{\beta}_{i,j}$ out of the objective function. This can be done by realizing that, conditional on the parameter vector $\boldsymbol{\theta}_{i,j}$, the $\hat{\boldsymbol{\beta}}_{i,j}$ that minimizes $Q_{i,j}$ equals

$$\hat{\boldsymbol{\beta}}_{i,j} = \left(\mathbf{X}_{i,j}'\mathbf{Z}_{i,j}\mathbf{W}_{i,j}\mathbf{Z}_{i,j}'\mathbf{X}_{i,j}\right)^{-1}\mathbf{X}_{i,j}'\mathbf{Z}_{i,j}\mathbf{W}_{i,j}\mathbf{Z}_{i,j}'\left(\mathbf{m}_{i,j} - \boldsymbol{\xi}_{i,j}\right) \tag{100}$$

Substituting this solution back into the objective function $Q_{i,j}$ yields the concentrated objective function

$$\widetilde{Q}_{i,j} = \frac{1}{T^2}\left(\mathbf{m}_{i,j} - \boldsymbol{\xi}_{i,j}\right)'\widetilde{\mathbf{W}}_{i,j}\left(\mathbf{m}_{i,j} - \boldsymbol{\xi}_{i,j}\right) \tag{101}$$

where

$$\begin{aligned} \widetilde{\mathbf{W}}_{i,j} = &\left[\mathbf{Z}_{i,j} - \mathbf{Z}_{i,j}\mathbf{W}_{i,j}\mathbf{Z}_{i,j}'\left(\mathbf{X}_{i,j}'\mathbf{Z}_{i,j}\mathbf{W}_{i,j}\mathbf{Z}_{i,j}'\mathbf{X}_{i,j}\right)^{-1}\mathbf{X}_{i,j}'\mathbf{Z}_{i,j}\right]\mathbf{W}_{i,j} \times \\ &\left[\mathbf{Z}_{i,j}' - \mathbf{Z}_{i,j}'\mathbf{X}_{i,j}\left(\mathbf{X}_{i,j}'\mathbf{Z}_{i,j}\mathbf{W}_{i,j}\mathbf{Z}_{i,j}'\mathbf{X}_{i,j}\right)^{-1}\mathbf{X}_{i,j}'\mathbf{Z}_{i,j}\mathbf{W}_{i,j}\mathbf{Z}_{i,j}'\right] \end{aligned} \tag{102}$$

such that $\widetilde{Q}_{i,j}$ only depends on the structural parameters that are to be estimated. That is, $\widetilde{Q}_{i,j}$ depends on $\boldsymbol{\theta}_{i,j}$ through $\boldsymbol{\xi}_{i,j}$. Consequently, the GMM estimate $\hat{\boldsymbol{\theta}}_{i,j}$ is obtained by minimizing $\widetilde{Q}_{i,j}$ with respect to $\boldsymbol{\theta}_{i,j}$.

Andrews and McDermott (1995) show that the standard errors of the GMM parameter estimates in this case can be calculated in the same way as under standard non-linear GMM estimation that does not contain any deterministic trends in the moment conditions.

If we define the gradient of the moment condition $\boldsymbol{\mu}_{i,j}$ as

$$\mathbf{D}_{\boldsymbol{\mu}_{ij}} = \begin{bmatrix} \dfrac{\partial \boldsymbol{\mu}_{i,j}}{\partial \boldsymbol{\theta}'_{i,j}} & \dfrac{\partial \boldsymbol{\mu}_{i,j}}{\partial \boldsymbol{\beta}'_{i,j}} \end{bmatrix} = -\frac{1}{T}\mathbf{Z}'_{i,j}\begin{bmatrix} \dfrac{\partial \boldsymbol{\xi}_{i,j}}{\partial \boldsymbol{\theta}'_{i,j}} & \mathbf{X}_{i,j} \end{bmatrix} \tag{103}$$

then the asymptotic covariance matrix of the parameter estimates $\begin{bmatrix} \hat{\boldsymbol{\theta}}'_{i,j} & \hat{\boldsymbol{\beta}}'_{i,j} \end{bmatrix}$ can be consistently estimated by

$$\hat{\mathbf{V}} = \left(\hat{\mathbf{D}}'_{\boldsymbol{\mu}_{ij}} \mathbf{W}_{i,j} \hat{\mathbf{D}}_{\boldsymbol{\mu}_{ij}}\right)^{-1} \hat{\mathbf{D}}'_{\boldsymbol{\mu}_{ij}} \mathbf{W}_{i,j} \hat{\mathbf{S}}_{i,j} \mathbf{W}_{i,j} \hat{\mathbf{D}}_{\boldsymbol{\mu}_{ij}} \left(\hat{\mathbf{D}}'_{\boldsymbol{\mu}_{ij}} \mathbf{W}_{i,j} \hat{\mathbf{D}}_{\boldsymbol{\mu}_{ij}}\right)^{-1} \tag{104}$$

where $\hat{\mathbf{S}}_{i,j}$ is a consistent estimator of $\mathbf{S}_{i,j} = TE\left[\boldsymbol{\mu}_{i,j}\boldsymbol{\mu}'_{i,j}\right]$. In this version of our estimation, $\hat{\mathbf{S}}_{i,j}$ is estimated under the assumption that $u_{i,K,j,t}(\tau) \sim IID(0,\sigma_{ij\tau}^2)$.

Table 1. Countries and technology measures covered in the data.

j	Country	i	τ	$v_{i\tau}$-v_{i1}	type	Technology group/vintage
1.	Australia	1				TEXTILES
2.	Austria		1	*0*	*K*	Number of mule spindles
3.	Belgium		2	*49*	*K*	Number of ring spindles
4.	Canada	2				STEEL
5.	Denmark		1	*0*	*Y*	Steel tonnage produced using Open Hearth furnaces
6.	Finland		2	*55*	*Y*	Steel tonnage produced using Blast Oxygen and Electric Arc furnaces
7.	France	3				SHIPPING
8.	Germany		1	-	*K*	Tonnage of sailships in merchant fleet
9.	Greece		2	*0*	*K*	Tonnage of steamships in merchant fleet
10.	Ireland		3	*96*	*K*	Tonnage of motorships in merchant fleet
11.	Italy	4				TELECOMMUNICATIONS
12.	Japan		1	*0*	*Y*	Telegrams sent
13.	Netherlands		2	*41*	*K*	Number of mainline telephones
14.	New Zealand		3	*112*	*K*	Number of mobile telephones
15.	Norway	5				MASS COMMUNICATIONS
16.	Portugal		1	*0*	*K*	Number of radios
17.	Spain		2	*30*	*K*	Number of televisions
18.	Sweden	6				TRANSPORTATION - CARGO
19.	Switzerland		1	*0*	*Y*	Freight traffic on railways (TKMs)
20.	United Kingdom		2	*51*	*K*	Number of commercial trucks
21.	United States		3	*78*	*Y*	Aviation cargo (TKMs)
		7				TRANSPORTATION – PASSENGERS
			1	*0*	*Y*	Passenger traffic on railways (PKMs)
			2	*51*	*K*	Number of passenger cars
			3	*78*	*Y*	Aviation passengers (PKMs)

Table 2. Parameter estimates

technology	Textiles		Steel		Shipping	
	value	std.err.	value	std.err.	value	std.err.
$\left(v_{i1} - v_{ij0}^{*}\right)$						
Australia			95.23	3.03	293.33	242.86
Austria	150.88	214.69	93.63	2.64	190.75	87.86
Belgium	77.18	2.42	92.29	3.09	142.32	2.46
Canada	74.00	4.67	95.05	7.41	308.33	171.39
Denmark			132.81	329.21	297.80	139.82
Finland	208.33	3690.66	103.44	3.60	158.40	18.36
France	84.27	1.13	91.43	1.39	155.03	19.13
Germany	68.05	0.69	99.48	1.31	147.67	1.19
Greece			93.85	13403.08		
Ireland			91.85	4289.95		
Italy	71.00	0.78	95.63	2.10	308.33	104.12
Japan	68.66	2.67	93.33	1.47	184.80	5984.64
Netherlands	85.34	1.11	87.16	17.91	305.83	355.15
New Zealand			132.03	364.89	154.94	4.63
Norway	99.58	269.38	97.21	6877.00	143.54	15.07
Portugal			98.55	1165.61		
Spain	157.60	168.61	97.02	2.41		
Sweden	121.70	300.82	99.63	2.32	181.11	28.77
Switzerland	72.89	1.36	76.59	245.77		
United Kingdom	98.05	6.66	94.69	1.61	191.05	16.30
United States	61.91	4.73	91.76	0.91	152.91	12.05
ψ_{ij}						
Australia			1.97	0.47	0.16	0.28
Austria	0.09	0.34	1.31	0.08	0.16	0.13
Belgium	1.02	0.02	1.16	0.09	0.37	0.09
Canada	1.01	0.07	2.71	2.00	0.07	0.08
Denmark			0.17	1.03	0.10	0.09
Finland	0.27	8.64	1.30	0.08	0.26	0.09
France	0.98	0.02	0.63	0.06	0.20	0.07
Germany	0.84	0.05	1.12	0.16	0.45	0.01
Greece			3.83	1194.14		
Ireland			3.83	175.45		
Italy	0.74	0.04	0.49	0.06	0.13	0.10
Japan	1.02	0.09	2.91	0.15	0.52	34.19
Netherlands	1.49	0.05	0.63	0.78	0.11	0.27
New Zealand			2.60	42.52	0.51	0.03
Norway	0.17	0.82	1.33	60.12	0.32	0.04
Portugal			3.00	453.38		
Spain	0.09	0.27	1.18	0.07		
Sweden	0.15	0.79	1.02	0.03	0.38	0.11
Switzerland	0.83	0.04	1.86	8.65		
United Kingdom	0.69	0.08	1.31	0.06	0.23	0.03
United States	1.32	0.06	3.53	0.15	0.26	0.06

Table 2 (continued). Parameter estimates

technology	Tele-communications		Mass communications		Transport - cargo		Transport – passengers	
	Value	std.err.	value	std.err.	value	std.err.	value	std.err.
$\left(v_{i1} - v^*_{ij0}\right)$								
Australia	121.41	0.76	242.75	28.20	212.58	12.12	149.90	14.25
Austria	149.29	3.58	145.32	1.21	312.50	94.82	230.25	17.41
Belgium	130.22	1.37	139.31	1.23	230.34	30.70	113.11	1.17
Canada	219.95	41.43	211.25	101.67	164.76	23.30	132.16	3.63
Denmark	124.02	0.56	126.96	3.56	159.91	4.65	164.31	9.55
Finland	132.76	2.55	189.59	9.02	146.34	2.50	140.45	3.16
France	136.00	1.76	163.23	8.87	144.44	8.72	165.63	4.19
Germany	129.86	1.86	155.13	25.01	155.85	3.89	143.78	3.46
Greece	168.51	11.12						
Ireland	131.55	4.14	193.76	35.62				
Italy	135.39	1.46	125.33	1.23	156.08	3.29	145.03	3.74
Japan	114.78	1.34	130.32	5.77	108.23	1.16	120.16	1.21
Netherlands	120.97	0.92	149.98	1.63	165.54	3.37	164.40	5.05
New Zealand	122.81	0.67	168.26	116.95	166.13	6.01	304.51	251.31
Norway	119.86	0.36	120.64	4.91	159.71	3.33	201.93	14.64
Portugal	121.22	0.67	115.90	5.24				
Spain	131.70	1.73	166.97	39.43				
Sweden	120.04	0.26	151.80	4.09	223.58	14.83	160.19	4.90
Switzerland	149.00	4.06	152.54	11.06	289.21	43.25	223.79	14.42
United Kingdom	130.98	1.72	309.12	136.30	147.68	2.91	159.58	8.88
United States	131.21	1.66	113.75	7.40	159.47	6.27	136.21	2.45
ψ_{ij}								
Australia	0.42	0.03	0.36	0.10	0.28	0.03	0.45	0.07
Austria	0.41	0.02	1.58	0.05	0.09	0.07	0.22	0.03
Belgium	0.54	0.01	1.83	0.07	0.24	0.06	0.64	0.01
Canada	0.05	0.04	0.08	0.15	0.37	0.08	0.54	0.02
Denmark	0.43	0.02	1.78	0.13	0.45	0.02	0.38	0.04
Finland	0.57	0.02	0.70	0.09	0.53	0.01	0.49	0.02
France	0.49	0.01	1.05	0.18	0.34	0.05	0.35	0.01
Germany	0.53	0.01	1.14	0.58	0.41	0.02	0.43	0.01
Greece	0.48	0.05						
Ireland	0.72	0.04	0.66	0.31				
Italy	0.34	0.03	2.28	0.11	0.53	0.02	0.42	0.02
Japan	0.67	0.01	2.63	0.28	0.84	0.02	0.62	0.01
Netherlands	0.55	0.02	1.38	0.05	0.44	0.02	0.38	0.02
New Zealand	0.55	0.01	1.00	2.10	0.27	0.10	0.13	0.23
Norway	0.55	0.02	2.76	0.14	0.45	0.02	0.28	0.04
Portugal	0.42	0.05	3.71	0.79				
Spain	0.65	0.02	1.01	0.72				
Sweden	0.57	0.00	1.24	0.07	0.25	0.03	0.40	0.02
Switzerland	0.28	0.02	1.18	0.27	0.14	0.05	0.22	0.03
United Kingdom	0.56	0.01	0.18	0.20	0.46	0.01	0.39	0.03
United States	0.46	0.01	3.40	1.47	0.41	0.03	0.51	0.01

Table 3. Estimated ranges of technologies in use (in years)

technology	Textiles		Steel		Shipping	
	$\underline{v}$	$\bar{v}$	$\underline{v}$	$\bar{v}$	$\underline{v}$	$\bar{v}$
Australia			-46	10	-281	-75
Austria	-68	181	-42	26	-179	25
Belgium	-80	-3	-39	33	-168	-27
Canada	-77	1	-47	2	-248	22
Denmark			-32	167	-257	-13
Finland	-190	-22	-51	17	-174	-1
France	-87	-8	-31	72	-156	33
Germany	-69	17	-46	27	-176	-49
Greece			-48	-3		
Ireland			-46	-1		
Italy	-70	24	-31	89	-287	-68
Japan	-72	6	-46	2	-216	-100
Netherlands	-92	-28	-26	76	-272	-39
New Zealand			-84	-35	-186	-69
Norway	-55	145	-45	22	-166	-10
Portugal			-52	-4		
Spain	-75	174	-44	27		
Sweden	-68	142	-45	32	-207	-69
Switzerland	-73	14	-27	30		
United Kingdom	-96	1	-43	25	-200	-20
United States	-67	0	-46	0	-170	0

technology	Tele-communications		Mass communications		Transport - cargo		Transport – passengers	
	$\underline{v}$	$\bar{v}$	$\underline{v}$	$\bar{v}$	$\underline{v}$	$\bar{v}$	$\underline{v}$	$\bar{v}$
Australia	-114	18	-151	-7	-179	-11	-129	-4
Austria	-141	-8	-73	-12	-217	29	-188	-6
Belgium	-127	-13	-68	-11	-190	-14	-98	3
Canada	-104	195	-42	219	-137	5	-115	-2
Denmark	-117	12	-56	2	-135	-10	-141	-2
Finland	-130	-21	-109	-13	-125	-11	-121	-1
France	-131	-10	-88	-12	-115	33	-140	7
Germany	-126	-12	-80	-8	-130	3	-122	8
Greece	-163	-42						
Ireland	-132	-38	-112	-13				
Italy	-123	27	-56	-3	-134	-20	-123	8
Japan	-115	-16	-61	-12	-93	-7	-105	-1
Netherlands	-118	-6	-77	-11	-141	-14	-141	-2
New Zealand	-119	-8	-92	-14	-132	36	-230	-8
Norway	-117	-6	-52	-3	-135	-10	-172	-7
Portugal	-113	18	-49	-4				
Spain	-131	-30	-91	-13				
Sweden	-117	-8	-78	-9	-185	-12	-137	-1
Switzerland	-133	34	-78	-7	-218	-6	-180	3
United Kingdom	-128	-18	-190	5	-123	2	-136	2
United States	-125	0	-46	0	-133	0	-117	0

Note: $\bar{v}$ *=0 for the USA is the normalization*

Table 4. Productivity estimates

technology	Textiles		Steel		Shipping	
ln(Z^p_ij)	Value	std.err.	value	std.err.	value	std.err.
Australia			-0.01	0.01	-1.82	2.46
Austria	-0.62	1.30	0.06	0.03	-0.39	0.86
Belgium	-0.18	0.03	0.09	0.04	0.06	0.23
Canada	-0.13	0.07	-0.03	0.00	-1.83	2.24
Denmark			-0.04	1.46	-1.76	1.52
Finland	-1.72	34.99	-0.08	0.03	-0.07	0.07
France	-0.27	0.01	0.21	0.01	0.05	0.08
Germany	-0.02	0.01	0.00	0.02	-0.06	0.01
Greece			-0.03	159.37		
Ireland			0.00	55.99		
Italy	-0.04	0.01	0.21	0.02	-1.99	1.08
Japan	-0.06	0.04	-0.01	0.01	-0.62	54.37
Netherlands	-0.34	0.01	0.27	0.24	-1.90	3.80
New Zealand			-0.55	2.92	-0.20	0.04
Norway	-0.06	1.22	0.00	85.51	0.08	0.14
Portugal			-0.09	1.47		
Spain	-0.72	1.01	0.02	0.03		
Sweden	-0.34	1.45	0.01	0.02	-0.50	0.26
Switzerland	-0.08	0.02	0.26	2.67		
United Kingdom	-0.40	0.08	0.04	0.02	-0.49	0.16
United States	0.00	0.06	0.00	0.00	0.00	0.05

technology	Tele-communications		Mass communications		Transport - cargo		Transport – passengers	
ln(Z^p_ij)	Value	std.err.	value	std.err.	value	std.err.	value	Std.err.
Australia	0.16	0.04	-1.47	0.26	-0.64	0.11	-0.17	0.13
Austria	-0.23	0.03	-0.38	0.01	-1.75	1.23	-1.10	0.17
Belgium	-0.02	0.02	-0.31	0.01	-0.85	0.29	0.27	0.01
Canada	-0.69	0.33	-0.63	0.64	-0.05	0.22	0.04	0.03
Denmark	0.11	0.03	-0.14	0.04	-0.03	0.04	-0.33	0.09
Finland	-0.07	0.04	-0.88	0.08	0.12	0.03	-0.05	0.03
France	-0.08	0.02	-0.58	0.08	0.26	0.20	-0.32	0.04
Germany	-0.02	0.02	-0.48	0.22	0.05	0.04	-0.07	0.04
Greece	-0.53	0.12						
Ireland	-0.11	0.04	-0.93	0.32				
Italy	0.02	0.05	-0.14	0.02	-0.01	0.03	-0.08	0.05
Japan	0.14	0.01	-0.22	0.07	0.56	0.02	0.18	0.01
Netherlands	0.10	0.02	-0.43	0.01	-0.11	0.03	-0.33	0.05
New Zealand	0.08	0.01	-0.65	1.01	0.01	0.18	-1.99	2.63
Norway	0.11	0.01	-0.08	0.07	-0.03	0.03	-0.77	0.14
Portugal	0.16	0.05	-0.03	0.06				
Spain	-0.09	0.03	-0.63	0.34				
Sweden	0.11	0.00	-0.45	0.05	-0.77	0.14	-0.28	0.04
Switzerland	-0.12	0.07	-0.45	0.10	-1.53	0.44	-1.00	0.14
United Kingdom	-0.05	0.02	-2.22	1.34	0.14	0.03	-0.26	0.08
United States	0.00	0.02	0.00	0.06	0.00	0.06	0.00	0.02

Note: Shaded estimates are the ones with a standard error smaller than 0.1.
Numbers are reported in deviation from U.S. level
Standard errors are computed taking as fixed the level of productivity for the US.

Table 5. Comparison with sectoral TFP levels

technology	Textiles		Steel		Shipping	
	ln(Z$^p_{ij}$)	*ln(Z$_{ij}$)*	*ln(Z$^p_{ij}$)*	*ln(Z$_{ij}$)*	*ln(Z$^p_{ij}$)*	*ln(Z$_{ij}$)*
N	10	11	13	11	5	13
std.dev.	0.14	0.59	0.09	0.75	0.09	0.74
N in common	8		9		4	
std.dev in common	0.14	0.64	0.09	0.84	0.06	0.08
correlation	0.36		-0.20		0.66	

technology	Tele-communications		Mass communications		Transport - cargo		Transport – passengers	
	ln(Z$^p_{ij}$)	*ln(Z$_{ij}$)*	*ln(Z$^p_{ij}$)*	*ln(Z$_{ij}$)*	*ln(Z$^p_{ij}$)*	*ln(Z$_{ij}$)*	*ln(Z$^p_{ij}$)*	*ln(Z$_{ij}$)*
N	19	6	13	6	9	13	12	13
std.dev.	0.11	1.14	0.25	1.14	0.20	0.74	0.20	0.74
N in common	6		5		8		11	
std.dev in common	0.09	1.14	0.21	1.26	0.20	0.93	0.20	0.74
correlation	0.00		0.23		0.83		0.65	

Note: Our measures of sectoral TFP levels, i.e. ln(Z$_{ij}$), are taken from Bernard and Jones (1996a).
Our technology groups are matched with the 1985 TFP levels of the following sectors from their data set respectively:
Textiles (TEX), Basic metal products (BMI), Transportation and communication (TRS), Communication (COM),
Communication (COM), Transportation and communication (TRS), and Transportation and communication (TRS).
Explanation: N total number countries for which observations, std.dev is based on all observations,
'N in common' is number of countries for which observations for both ln(Z$^p_{ij}$) and ln(Z$_{ij}$),
'std.dev. in common' is based on the common observations, so is the correlation.

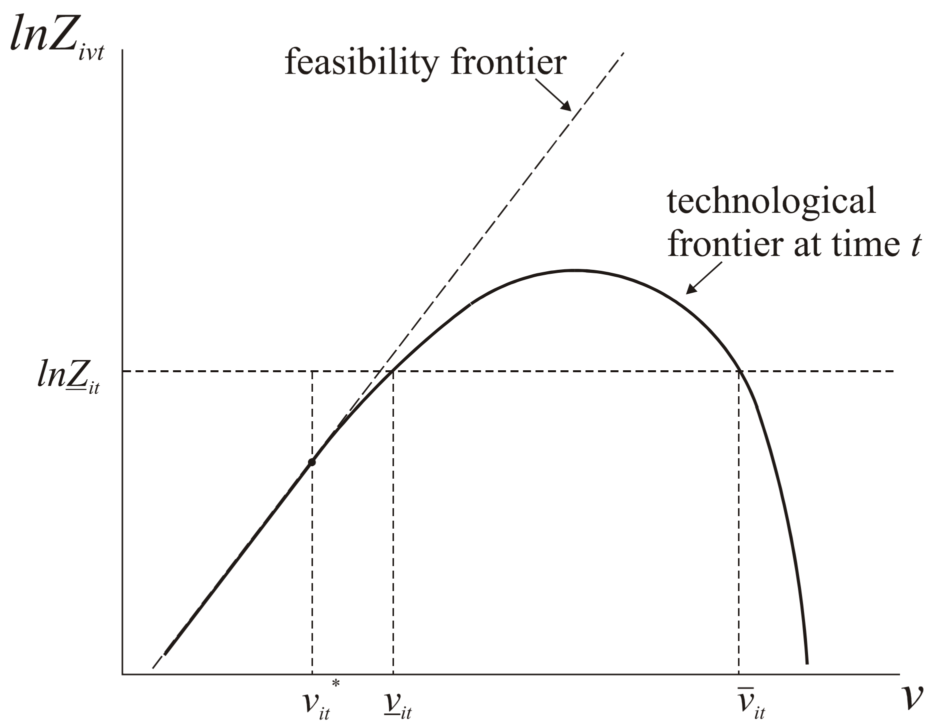

Figure 1. Shape of the vintage technology frontier

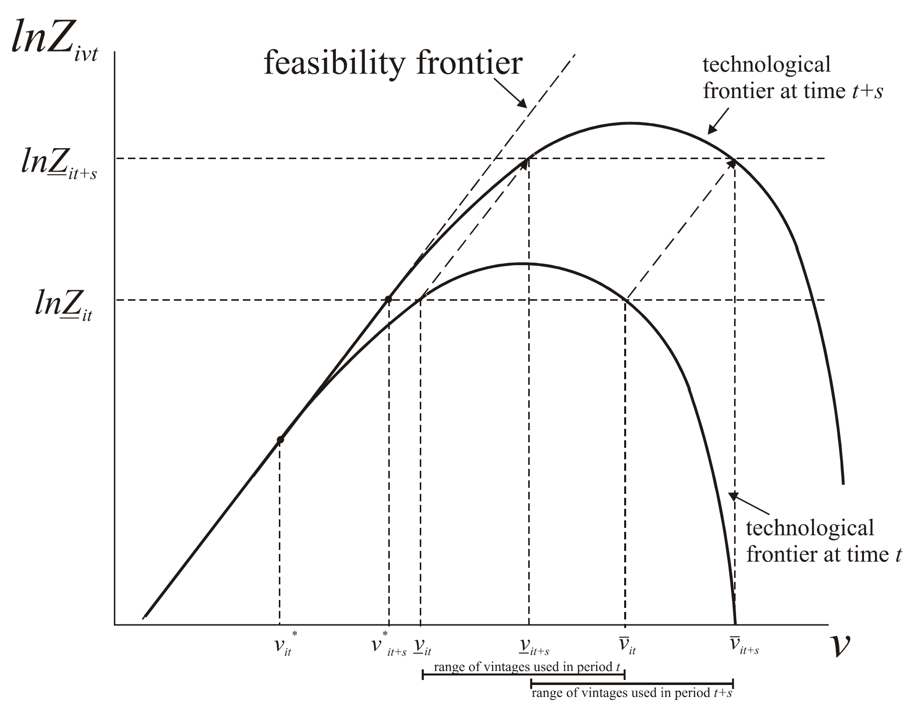

Figure 2. Sectoral technological progress

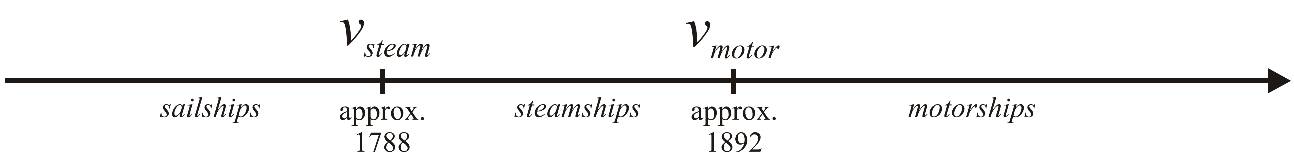

Figure 3. discretization of continuum of vintages for merchant shipping

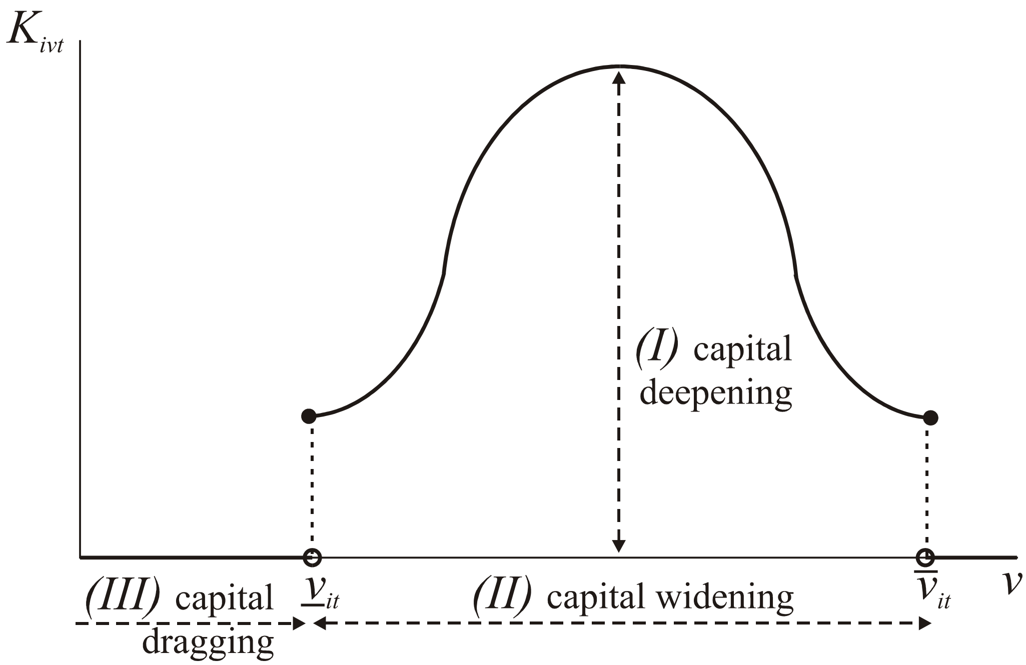

Figure 4. Capital deepening versus capital widening and dragging

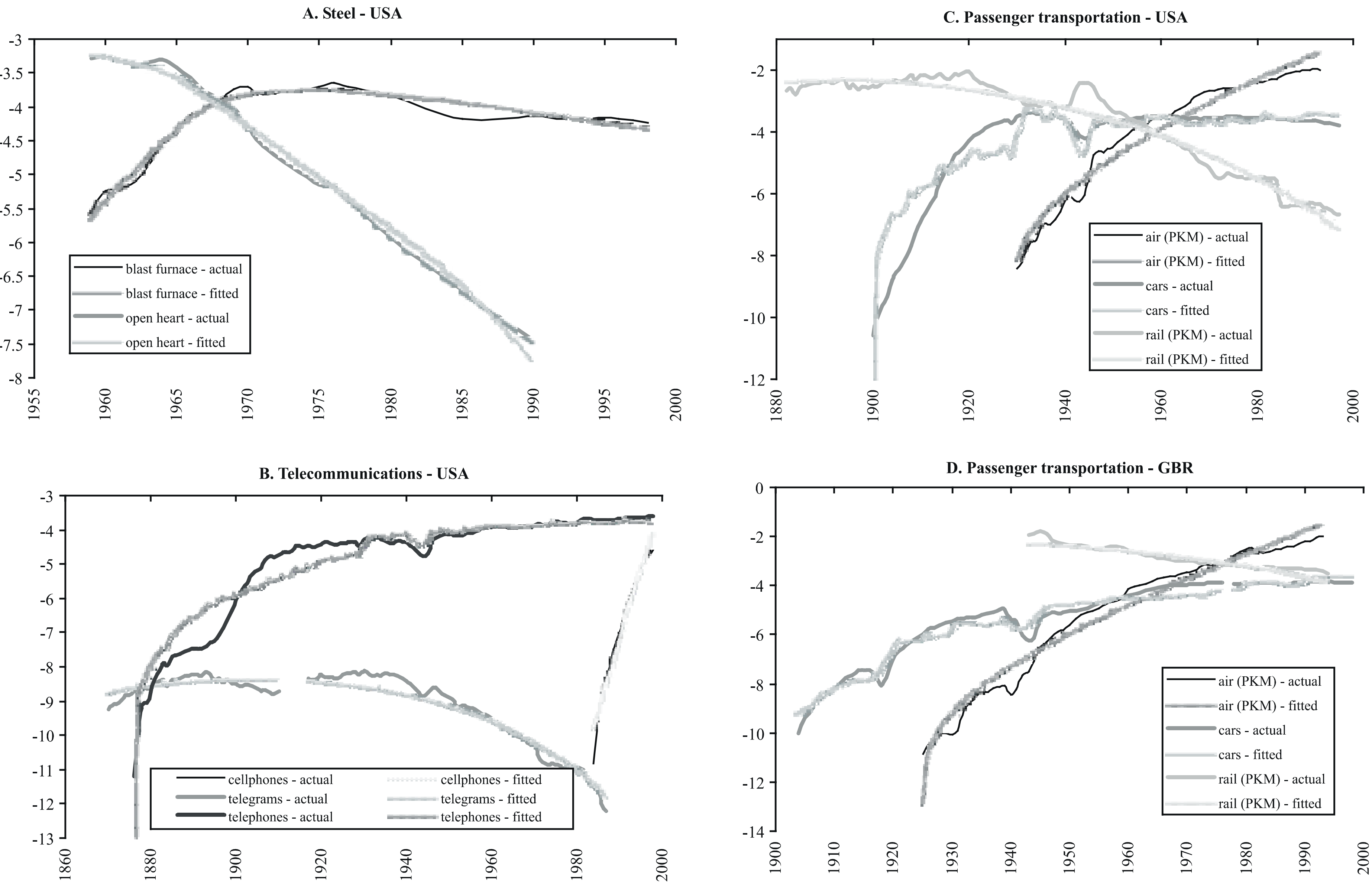

Figure 5. Fit of adoption curves for three technologies for the U.S. and one for the U.K.