DISCUSSION PAPER SERIES

No. 3755

FACTOR ENDOWMENTS AND PRODUCTION IN EUROPEAN REGIONS

Stephen Redding and Mercedes Vera-Martin

INTERNATIONAL TRADE

ISSN 0265-8003

FACTOR ENDOWMENTS AND PRODUCTION IN EUROPEAN REGIONS

Stephen Redding, London School of Economics (LSE) and CEPR
Mercedes Vera-Martin, London School of Economics (LSE)

Discussion Paper No. 3755
February 2003

Centre for Economic Policy Research
90–98 Goswell Rd, London EC1V 7RR, UK
Tel: (44 20) 7878 2900, Fax: (44 20) 7878 2999
Email: cepr@cepr.org, Website: www.cepr.org

CEPR Discussion Paper No. 3755

February 2003

ABSTRACT

Factor Endowments and Production in European Regions*

This Paper analyses patterns of production across 14 industries in 45 regions from seven European countries since 1975. We estimate a structural equation derived directly from neoclassical trade theory that relates an industry's share of a region's GDP to factor endowments, relative prices and technology. Although factor endowments play a statistically significant and quantitatively important role in explaining production patterns, the Heckscher-Ohlin model is rejected against more general alternatives that allow for regional variation in relative prices and technology. Factor endowments are more successful at explaining patterns of production in aggregate industries (Agriculture, Manufacturing, Services) than in disaggregated industries within Manufacturing.

JEL Classification: F11, F14 and R13
Keywords: european integration, factor endowments, Heckscher-Ohlin, neoclassical model and specialization

Stephen Redding
London School of Economics
Houghton Street
LONDON
WC2A 2EA
Tel: (44 20) 7955 7483
Fax: (44 20) 7831 1840
Email: s.j.redding@lse.ac.uk

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=135276

Mercedes Vera-Martin
London School of Economics
Houghton Street
LONDON
WC2A 2AE
Tel: (44 20) 7955 7800
Fax: (44 20) 7955 7585
Email: m.vera-martin@lse.ac.uk

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=144263

*This Paper is produced as part of the Globalisation Programme of the ESRC-funded Centre for Economic Performance and is a revised version of CEP Discussion Paper 501. A Central Bank of Spain Scholarship funded Vera-Martin's research. We are very grateful to Oriana Bandiera, James Harrigan, Steve Nickell, Francois Ortalo-Magne, Henry Overman, Peter Schott, Jo Swaffield, Tony Venables, Alan Winters, and seminar participants at the London School of Economics and Yale University for helpful comments and suggestions. We would like to thank Maia Guell-Rotllan, Steve Machin, Marco Manacorda, Steve Nickell, Jan van Ours, Henry Overman, and Jo Swaffield for their help with the data. We are also grateful to Giorgia Albertin and Estela Montado for research assistance. Responsibility for any results, opinions, and errors lies with the authors alone.

Submitted 22 January 2002

1 Introduction

> "One of the best ways to understand how the international economy works is to start looking at what happens inside nations ... The data will be better and pose fewer problems of compatibility, and the underlying economic forces will be less distorted by government policies."[1]

One of the most influential conceptual frameworks for theoretical and empirical work in international trade is the Heckscher-Ohlin (HO) model. A key attraction is the model's ability to yield precisely formulated theoretical predictions which are amenable to direct empirical testing. However, a number of cross-country studies have called into question its empirical validity. For example, using data on cross-country trade in factor services, Bowen *et al.* (1987), Trefler (1995) and Davis and Weinstein (2001) reject the HO model's assumptions of identical and homothetic preferences, identical technologies, and no barriers to trade against a variety of more general alternatives.[2] Similarly, when examining the cross-country relationship between industry-level production and factor endowments, Harrigan (1995) finds large within-sample prediction errors, while Harrigan (1997) provides evidence that non-neutral technology differences play an important role in explaining cross-country variation in production structure.

This paper examines the ability of the HO model to explain production patterns at the regional level in Europe using a newly constructed panel dataset on output in 14 industries and endowments of 5 factors of production for 45 NUTS-1 regions from 7 European countries since 1975.[3] The use of regional data enables us to abstract from many of the reasons advanced for the poor performance of the HO model at the country-level. For example, both measurement error and technology differences are likely to be much smaller across regions within Europe than for a cross-section of developed and developing countries. The ongoing process of economic integration within the European Union means that it is an interesting context within which to explore the relationship between production and factor endowments. We control for exogenous variation in relative prices induced by European integration and examine whether this process of international integration has strengthened or weakened the relationship between production and factor endowments across regions within countries.

Much existing empirical work on the international location of production has, for reasons of data availability, been concerned with the manufacturing sector. This paper explicitly considers both manufacturing and non-manufacturing, where the latter accounts for more than 70% of GDP in many

[1] Paul Krugman, *Geography and Trade*, MIT Press, 1991, page 3, cited in Bernstein and Weinstein (2002).

[2] These more general alternatives allow, among other things, for cross-country differences in technology, non-factor price equalization, trade costs and measurement error. See also Davis *et al.* (1997) and Gabaix (1997).

[3] NUTS stands for Nomenclature of Statistical Territorial Units. NUTS-1 regions are the first-tier of sub-national geographical units for which Eurostat collects data on the EU member countries. See Appendix A for more details concerning the data used.

NUTS-1 regions. We analyze production structure at two alternative levels of industrial aggregation. First, we consider the 3 aggregate industries of Agriculture, Manufacturing and Services. Second, we break out Manufacturing into 11 more disaggregated industries. We consider endowments of 5 factors of production: high-education, medium-education, and low education individuals, physical capital, and land area. The analysis focuses on patterns of production rather than trade, because the central predictions of the HO model are for producer equilibrium and, in so doing, we abstract from any violations of the model's assumptions concerning consumer behaviour.[4]

The paper derives a general equilibrium relationship between production structure and factor endowments that holds under the null hypothesis of the HO model with its assumptions of identical prices and technologies. We compare this with the relationship that holds under the more general alternative hypothesis of the neoclassical model of trade which allows for regional variation in both relative prices and technology. We are able to explicitly test for European regions whether the HO null is rejected against neoclassical alternatives, and use our framework to examine the quantitative importance of factor endowments relative to other considerations in explaining regional variation in production patterns.

While the use of regional data has many advantages, it means that the standard trade assumption that endowments of factors of production are exogenous and perfectly immobile across locations is less likely to apply. We show that the general equilibrium relationship between production structure and factor endowments under the null hypothesis and the corresponding relationship under the alternative hypothesis hold irrespective of whether factors of production are perfectly immobile or perfectly mobile across locations [5]

Factor mobility does, however, change the interpretation of these relationships. If factor endowments are exogenous and perfectly immobile across locations, the general equilibrium relationship between production structure and factor endowments has a *supply-side* interpretation. Changes in factor endowments cause changes in production structure (production moves in response to factor endowments). If factor endowments are mobile across locations, they become potentially endogenous to production structure. In addition to the *supply-side* interpretation given above, there is also a *demand-side* interpretation whereby changes in production structure cause factor endowments to move across regions (factor endowments move in response to production structure). Irrespective of

[4] Three of the HO model's four key theorems - the Rybczynski, Stolper-Samuelson, and Factor Price Equalization Theorems - require no assumptions about consumer preferences.

[5] In practice, a wide range of evidence suggests that factor mobility across European regions is relatively low. This is particularly true across countries, where language and cultural differences act as barriers to labour mobility. However even within European countries, there is evidence that labour mobility is relatively low: see for example McCormick (1997) and Cameron and Muellbauer (1999) for evidence on the UK.

whether the relationships we estimate are demand-side, supply-side or a combination of both, we are able to test the HO model's predictions for the relationship between production and endowments against those of the more general neoclassical model. Similarly, irrespective of which interpretation applies, we can take the more general neoclassical model and examine the respective contributions of factor endowments and other considerations in statistically explaining variation in production structure across European regions.

Our main empirical findings are as follows. First, the HO model provides an incomplete explanation of patterns of production across European regions and is rejected against more general neoclassical alternatives. Second, although the HO model is rejected, factor endowments remain statistically significant and quantitatively important in explaining production structure within these neoclassical alternatives. Individual factor endowments are highly statistically significant and including information on factor endowments reduces the model's within-sample average absolute prediction error by a factor of around 3 in Manufacturing.

Third, the pattern of estimated coefficients on factor endowments across industries is generally consistent with economic priors regarding factor intensity. For example, physical capital endowments are positively correlated with the share of Manufacturing in GDP and negatively correlated with the shares of Agriculture and Services. Higher numbers of medium education individuals relative to low education individuals are associated with a lower share of Agriculture in GDP and a higher share of Manufacturing. Higher numbers of high education individuals relative to medium education individuals are associated with a lower share of Manufacturing in GDP and a higher share of Services.

Fourth, factor endowments are more successful in explaining patterns of production at the aggregate level in Agricultural, Manufacturing and Services (where we have 3 industries and either 3 or 5 factor endowments) than in disaggregated manufacturing industries (where we have 11 industries and either 3 or 5 factor endowments). Within-sample average absolute prediction errors are typically far larger in the disaggregated manufacturing industries, and this is exactly as theory would predict. In the HO model with identical prices and technology and with no joint production, patterns of production are only determinate if there are at least as many factors of production as goods. Therefore, production indeterminacy provides one explanation for larger average absolute prediction errors in disaggregated manufacturing industries. Another explanation, again consistent with the theory, is that regional price and technology differences not controlled for in the right-hand side variables are particularly large in individual manufacturing industries.

Fifth, we find no evidence that the process of increasing economic integration in Europe has weakened the relationship between patterns of production and factor endowments across regions within countries. Our baseline econometric specification includes country-year dummies so that the coefficients on factor endowments' are identified solely from variation across regions within countries. Examining within-sample prediction errors for this specification reveals no systematic trend over time for either the 3 aggregate industries or the 11 disaggregated industries within manufacturing.

The paper is related to four main strands of existing work. First, as discussed above, there are a number of studies which have examined the relationship between factor endowments and production at the country-level, including Harrigan (1995, 1997), Harrigan and Zakrajsek (2000), Nickell *at al.* (2001), Redding (2002), and Schott (2002).[6] Second, there is an emerging empirical literature which has recently begun to examine the predictions of the HO model using regional-data. Thus, Davis *et al.* (1997) provide evidence that the HO model is more successful in explaining net trade in factor services between Japanese regions than between countries. Bernstein and Weinstein (2002) examine the relationship between production and factor endowments across Japanese prefectures. Although the data are consistent with factor price equalization, they find substantial within-sample prediction errors which are interpreted as evidence of production indeterminacy. Hanson and Slaughter (2002) use data on immigration for US states to test a generalization of the Rybczynski Theorem which predicts that regions will accommodate immigrant inflows by changes in output mix rather than changes in relative factor prices.[7] Changes in state output mix broadly match changes in state endowments, and the variation in factor intensities across states is found to be consistent with the equalization of relative factor prices.

Third, there is a body of more descriptive work which has sought to characterize patterns of specialization in Europe and their evolution over time. The vast majority of this work is undertaken at the country-level data (see, for example, Amiti 1999, Brulhart 2000, Midelfart-Knarvik, Overman, Redding and Venables 2000, and Proudman and Redding 1998, 2000), although two exceptions using regional data are Molle (1996) and Vera-Martin (2002).

Fourth, a smaller number of studies have examined the role of economic geography in shaping the location of production (see Overman *et al.* 2002 for a survey of this literature). Thus, Davis and Weinstein (1999), (2002) use data on a cross-section of countries and on Japanese regions to test for the 'home market' or 'magnification' effect, where an increase in expenditure on a good leads to a more than proportionate increase in production. Midelfart-Knarvik, Overman and Venables (2000) examine the roles of traditional trade theory (eg factor endowments and intensities) and economic geography (eg forward and backward linkages) in determining the cross-country location

[6] As also discusssed above, a related literature considers the cross-country relationship between international trade in factor services and factor endowments. See, for example, Leamer (1984), Bowen *et al.* (1987), Trefler (1995), Davis *et al.* (1997), Gabaix (1997) and Davis and Weinstein (2001).

[7] See Gandal, Hanson and Slaughter (2002) for a related analysis of immigration in Israel.

of manufacturing production.

The remainder of the paper is organized as follows. Section 2 introduces the theoretical framework and derives predictions for production patterns under the HO null and the neoclassical alternative. Section 3 describes the European regional production and factor endowments data. Section 4 discusses the econometric specification. Section 5 presents the estimation results. Section 6 concludes.

2 Theoretical Framework

The theoretical framework is provided by the neoclassical theory of trade and production (see, in particular, the exposition in Dixit and Norman 1980). Regions are indexed by $z \in \{1, ..., Z\}$; goods by $j \in \{1, ..., N\}$; factors of production by $i \in \{1, ..., M\}$; and time by t. Production is assumed to occur under conditions of perfect competition and constant returns to scale.[8] The neoclassical model allows for regional differences in factor endowments as well as region-industry differences in technology and relative prices. The HO model corresponds to a special case where all regions have identical relative prices and technology, and is therefore nested by the neoclassical model.

General equilibrium in production may be represented using the revenue function $r_z(p_{zt}, v_{zt})$, where p_{zt} denotes a region's vector of relative prices and v_{zt} is its vector of factor endowments. Under the assumption that the revenue function is twice continuously differentiable, we obtain determinate predictions for a region's vector of profit-maximizing net outputs $y_z(p_{zt}, v_{zt})$ which equals the gradient of $r_z(p_{zt}, v_{zt})$ with respect to p_{zt}.[9] We allow for Hicks-neutral region-industry-time technology differences so that the production technology takes the form $y_{zjt} = \theta_{zjt} F_j(v_{zjt})$, where θ_{zjt} parameterizes technology or productivity in industry j of region z at time t.[10] In this case, the revenue function function takes the form $r_z(p_{zt}, v_{zt}) = r(\theta_{zt} p_{zt}, v_{zt})$, where θ_{zt} is an $N \times N$ diagonal matrix of the technology parameters θ_{zjt}.[11] Changes in technology in industry j of region z have analogous effects on revenue to changes in industry j prices.

[8] While analysis of the neclassical model typically focuses on the perfectly competitive case, it is also possible to analyze imperfect competition as discussed in Helpman (1984).

[9] A sufficient condition for the revenue function to be twice continuously differentiable and production patterns to be determinate is that there are at least as many factors as goods: $M \geq N$. In the HO model where relative prices and technology are identical, production levels may still be determinant when $N > M$ if there is joint production. More generally in the neoclassical model, differences in technology and relative prices may render production determinant when $N > M$. The potential existence of production indeterminacy is really an empirical issue which we investigate below for alternative numbers of goods and factors. If production indeterminacy exists, the equation that we derive under the null linking production and factor endowments will be relatively unsuccessful in explaining regions' production patterns, in terms of having statistically insignificant right-hand side variables, low explanatory power and large within-sample prediction errors.

[10] The technology differences may vary across industries but are Hicks-neutral in the sense that they raise the productivity of all factors of production in industry j of region z by the same proportion. It is also possible to examine factor augmenting technology differences, as discussed further in Dixit and Norman (1980).

[11] See Dixit and Norman (1980), pages 137-9.

We follow Harrigan (1997) and Kohli (1991) in assuming a translog revenue function. This flexible functional form provides an arbitrarily close local approximation to the true underlying revenue function:

$$\begin{aligned}\ln r\left(\theta_{zt}p_{zt}, v_{zt}\right) = \beta_{00} + \textstyle\sum_j \beta_{0j}\ln\theta_{zjt}p_{zjt} + \frac{1}{2}\sum_j\sum_k \beta_{jk}\ln(\theta_{zjt}p_{zjt})\ln(\theta_{zkt}p_{zkt}) \\ + \textstyle\sum_i \delta_{0i}\ln v_{zit} + \frac{1}{2}\sum_i\sum_h \delta_{ih}\ln v_{zit}\ln v_{zht} \\ + \textstyle\sum_j\sum_i \gamma_{ji}\ln(\theta_{zjt}p_{zjt})\ln(v_{zit})\end{aligned} \quad (1)$$

where $j, k \in \{1, .., N\}$ index goods and $i, h \in \{1, .., M\}$ index factors. Symmetry of the cross effects implies: $\beta_{jk} = \beta_{kj}$ and $\delta_{ih} = \delta_{hi}$ for all j, k, i, h. Linear homogeneity of degree 1 in v and p requires: $\sum_j \beta_{0j} = 1$, $\sum_i \delta_{0i} = 1$, $\sum_j \beta_{jk} = 0$, $\sum_i \delta_{ih} = 0$, and $\sum_i \gamma_{ji} = 0$. Differentiating the revenue function with respect to p_j, we obtain the following equation for the share of industry j in region z's GDP at time t:

$$s_{zjt} \equiv \frac{p_{zjt}y_{zjt}(p_{zt}, v_{zt})}{r\left(p_{zt}, v_{zt}\right)} = \beta_{0j} + \sum_k \beta_{jk}\ln p_{zkt} + \sum_k \beta_{jk}\ln\theta_{zkt} + \sum_i \gamma_{ji}\ln v_{zit} \quad (2)$$

Thus, the share of an industry in GDP (s_{zjt}) provides a natural and theory-consistent measure of a region's extent of specialization in an industry. Under the assumptions of the neoclassical model, this theory-consistent measure is related in general equilibrium to the region's vectors of relative prices, technology levels, and factor endowments according to equation (2). The translog specification implies coefficients on these variables that are constant across regions and over time. This is true even without factor price equalization and, with regional differences in prices and technology, factor price equalization will typically not be observed. The effect of regional differences in relative prices and technology on patterns of production is directly controlled for by the presence of the second and third terms on the right-hand side of the equation.

The analysis so far makes no assumptions about whether regions are large or small, and allows for both tradeable and non-tradeable goods. If regions are small and all goods are tradeable, relative prices will be exogenously determined on world markets. More generally, relative prices will themselves be endogenous. Factors of production may be either perfectly immobile or exhibit a degree of mobility across regions. In either case, the relationship in equation (2) must hold in general equilibrium.

Under the assumptions of the HO model, relative prices and technology are identical across regions. In this case, the terms for relative prices and technology on the right hand-side of equation (2) may be replaced by a set of time dummies. These time dummies have industry-specific coefficients (the coefficients β_{jk} vary across industries j), reflecting the fact that changes in relative prices and technology have different effects in different industries. Substituting for relative prices and technology, we obtain our *null hypothesis*:

$$\textbf{(NULL)} \qquad s_{zjt} = \sum_t \phi_{jt} d_t + \sum_i \gamma_{ji} \ln v_{zit} + \varepsilon_{zjt} \tag{3}$$

where d_t are $\{0, 1\}$ dummies for time periods; ϕ_{jt} are the industry-specific coefficients on the time dummies; ε_{zjt} is a stochastic error; and the constant β_{0j} from equation (2) has been absorbed in the industry-specific coefficients on the time dummies. Since all coefficients in equation (3) vary across industries j, this relationship may be estimated separately for each industry, pooling observations across regions and over time.

Under the *alternative hypothesis* of the neoclassical model, relative prices and technology may vary across regions, industries and time. Unfortunately, region-industry-time specific data on prices are not available for European regions, and it is not therefore possible to construct direct measures of relative prices and technical efficiency. Therefore, we follow Harrigan (1997) in modelling relative prices and technology as being drawn from an estimable probability distribution. We consider a series of progressively more general models of relative prices and technology, each of which when substituted in equation (2) provides a progressively more general alternative to the HO null. First, we model differences in relative prices and technology with a country-industry fixed effect (η_{cj}), industry-time dummies ($\mu_{jt} d_t$), and a stochastic error (u_{zjt}):

$$\sum_k \beta_{jk} \ln p_{zkt} + \sum_k \beta_{jk} \ln \theta_{zkt} = \eta_{cj} + \sum_t \mu_{jt} d_t + u_{zjt} \tag{4}$$

which yields our first alternative hypothesis:

$$\textbf{(ALT1)} \qquad s_{zjt} = \eta_{cj} + \sum_t \zeta_{jt} d_t + \sum_i \gamma_{ji} \ln v_{zit} + \omega_{zjt} \tag{5}$$

where the constant β_{0j} has again been absorbed in other coefficients. This specification differs from the null hypothesis through the inclusion of the country-industry fixed effect which allows for permanent cross-country differences in relative prices and technology that are non-neutral across industries.

Second, we generalize the model of relative prices and technology to allow for country-specific trends in relative prices and technology over time. We capture these by including country-time dummies, which again have industry-specific coefficients reflecting the fact that changes in relative prices impact differentially across industries:

$$\sum_k \beta_{jk} \ln p_{zkt} + \sum_k \beta_{jk} \ln \theta_{zkt} = \sum_c \sum_t \mu_{cjt} d_{ct} + u_{zjt} \tag{6}$$

where d_{ct} are $\{0,1\}$ country-time dummies and μ_{cjt} are the industry-specific coefficients on these country-time dummies. This yields our second alternative hypothesis:

$$\textbf{(ALT2)} \qquad s_{zjt} = \sum_c \sum_t \zeta_{cjt} d_{ct} + \sum_i \gamma_{ji} \ln v_{zit} + \omega_{zjt} \tag{7}$$

where the constant β_{0j} has again been absorbed in other coefficients, and this specification differs from the null hypothesis because the effects of the time dummies now vary across both countries and industries.

Equation **(ALT2)** allows for cross-country differences in relative prices and technology that are both non-neutral across industries and time-varying. It is consistent with empirical evidence from the literature on productivity measurement, which typically finds cross-country productivity differences with these properties.[12] It is also substantially more general than many existing studies in the empirical trade literature, which often focus on technology differences that are neutral across industries, and it allows European integration to have different effects on relative prices across countries.

Third, we extend the model of relative prices and technology further to allow for permanent region-industry specific differences in relative prices and technology (η_{zj}), country-specific trends in relative prices and technology over time (d_{ct}) and a stochastic error (u_{zjt}):

$$\sum_k \beta_{jk} \ln p_{zkt} + \sum_k \beta_{jk} \ln \theta_{zkt} = \eta_{zj} + \sum_c \sum_t \mu_{cjt} d_{ct} + u_{zjt} \tag{8}$$

which yields our third alternative hypothesis:

$$\textbf{(ALT3)} \qquad s_{zjt} = \eta_{zj} + \sum_c \sum_t \zeta_{cjt} d_{ct} + \sum_i \gamma_{ji} \ln v_{zit} + \omega_{zjt} \tag{9}$$

where the constant β_{0j} has again been absorbed in other coefficients. This specification differs from the null hypothesis, because of both coefficients on the time dummies that vary across countries and industries, and because of the inclusion of a region-industry fixed effect.

The null hypothesis is derived directly from the HO model with its assumption of identical relative prices and technology. Similarly, each of the alternative hypotheses is derived directly from the neoclassical model and involves making progressively more general assumptions about relative prices and technology. Both variable choice and functional form are shaped by the underlying theory. Hence, under the assumptions of the null hypothesis or a particular alternative hypothesis, the relevant relationship may be given a structural economic interpretation.

In particular, *under the assumptions of the null hypothesis*, the coefficients on factor endowments (γ_{ji}) in **(NULL)** are directly related to the *Rybczynksi derivatives* of HO theory - the general

[12] See, for example, Bernard and Jones (1996) and Griffith *et al.* (2000).

equilibrium relationship between production and factor endowments holding constant relative prices and technology.[13] As discussed above, if factor endowments exhibit a degree of mobility across regions, the Rybczynski derivatives capture a relationship between output and factor endowments that must hold in general equilibrium, but that may be given either a demand-side or supply-side interpretation. In relating the coefficients γ_{ji} to the Rybczynksi derivatives, we are *not* asserting the existence of a causal relationship between exogenous changes in factor endowments and endogenous changes in production, but are instead examining a relationship that holds in equilibrium (examining the equilibrium value of the derivatives).

The factor intensities of industries are directly captured in our analysis by the estimated coefficients (γ_{ji}) on factor endowments. With large numbers of factors (M) and goods (N), many conventional definitions of factor intensity (such as the ratio of use of one factor of production to another) are problematic. Nevertheless, the fact that in our approach the estimated coefficients on factor endowments are directly linked to the Rybczynski derivatives of HO theory means that they can be related to natural measures of factor intensity.

For example, with $M = N$, a positive value of a Rybczynski derivative for factor i and good j implies that, if the price of factor i increases by one unit and all other factor prices are adjusted to keep other goods' unit costs unchanged, the unit costs of production for good j must rise. In this sense, good j is intensive in the use of factor i relative to the economy as a whole when $\partial^2 r(p,v)/\partial p_j \partial v_i$ and hence γ_{ji} is positive . With $M \geq N$, there will generally be more than one set of values for other factor prices that leave other goods' unit costs unchanged. Nonetheless, a natural measure of factor intensity still exists based on the Rybczynski derivatives themselves, whereby good j is said to be relatively more intensive in factor i than the average if $\partial^2 r(p,v)/\partial p_j \partial v_i$ and hence γ_{ji} is positive.[14]

Intuitively, we do not try to construct conventional measures of factor intensity which are often problematic with large numbers of factors and goods, but instead directly estimate the relationship between production and factor endowments. Under the HO null, factor intensities are captured in the estimated coefficients on factor endowments. The fact that natural measures of factor intensity are defined relative to the whole structure of general equilibrium when there are large numbers of factors (M) and goods (N) is directly related to the way in which the theorems of the 2×2×2 HO model now hold in a weakened form as averages and correlations.[15]

Note that the estimated coefficients on factor endowments (γ_{ji}) in **(NULL)** only correspond to

[13] Formally, the *Rybczynski derivatives* are $\partial y_j(p,v)/\partial v_i = \partial^2 r(p,v)/\partial p_j \partial v_i = \partial^2 r(p,v)/\partial v_i \partial p_j$ for all j, i. Differentiating with respect to v_i in equation **(NULL)** and rearranging, $\partial^2 \ln(r(p,v))/\partial p_j \partial v_i = \gamma_{ji}/(p_j v_i)$ so that the *Rybczynski derivatives* take the same sign as the γ_{ji}.

[14] See Dixit and Norman (1980), pages 53-9.

[15] See, for example, Dixit and Norman (1980), Chapter 4, and Ethier (1984), Sections 6 & 7.

Rybczynski derivatives if the assumptions of the HO null are satisfied. *Under the alternative hypothesis* of the more general neoclassical model, **(NULL)** is mis-specified because it omits terms capturing regional variation in relative prices and technology. In general, these terms will be correlated with factor endowments and their omission will give rise to omitted variables bias and inconsistent estimates of the γ_{ji}. Only by including these additional terms, as for example in **(ALT2)**, can consistent estimates of the γ_{ji} be obtained.

We investigate the importance of regional variation in relative prices and technology by comparing the results of estimating **(NULL)** to those from estimating the three alternative specifications **(ALT1)-(ALT3)**. We compare the specifications along four main dimensions. First, since the null is nested by each of the alternatives, a test of the statistical significance of the additional terms capturing regional variation in relative prices and technology provides a formal test of the null against each of the more general alternatives. Second, we investigate the importance of the omitted variables bias by considering how the estimated coefficients on factor endowments change as we move to progressively more general models of relative prices and technology. Third, we examine the quantitative importance of factor endowments relative to other considerations in explaining regional variation in production structure. Fourth, we evaluate the empirical performance of the null and alternative hypotheses using a number of model specification tests.

3 Data Description

The main source of data is the Regio dataset compiled by the European Statistics Office (Eurostat). We analyze patterns of production across 14 industries in 45 NUTS-1 regions from 7 European countries since 1975. The choice of countries reflects the availability of data; we consider Belgium, France, Italy, Luxembourg, Netherlands, Spain and the United Kingdom.[16] As will be shown below, this is a group of countries among which there is substantial heterogeneity in patterns of production and factor endowments. The group includes several countries close to the 'core' of Europe (eg Belgium and France) and others located further towards the 'periphery' (eg Italy and Spain).

The number and size of NUTS-1 regions varies across European countries. This is perfectly consistent with our model, and the variation in size will be exploited in tests of the linear homogeneity restrictions implied by theory. In some European countries, such as Italy, the NUTS-1 regions correspond to the main regional political units. In the UK, they comprise geographical areas such as the North, South East, and South West. A full list of NUTS-1 regions in each country is given

[16]The data for other European countries are very incomplete. Where information is available, it is for a very short period of time.

in Appendix A. We show below that there is also substantial variation in specialization and factor endowments across NUTS-1 regions within a country - from, for example, the North of Italy to Sicily.

Patterns of production are analyzed at two alternative levels of aggregation. First, we consider three aggregate (one-digit) industries: Agriculture, Manufacturing and Services. Second, we exploit more disaggregated information on individual industries within Manufacturing. These are mainly two-digit industries and include, for example, Textiles & Clothing and Chemicals. Again, full details are given in Appendix A.

The Regio dataset provides information on industry value-added and GDP by region, from which we compute the share of each sector in GDP. It also provides information on three broad factor endowments: total population, physical capital and land area.[17] These data are merged with information on educational attainment at the regional level from individual country labour force surveys. This enables us to disaggregate the population endowment into low, medium and high education. The definitions we employ are standard in the labour market literature (see, for example, Nickell and Bell 1996 and Machin and Van Reenen 1998). 'Low education' corresponds to no or primary qualifications, 'medium education' denotes secondary and/or vocational qualifications, and 'high education' is college degree or equivalent.[18]

The length of the time-series available varies with the level of industrial aggregation, whether or not we use the information on educational attainment, and with the country considered. In order to exploit all of the information available, we consider two estimation samples. First, at the level of the three aggregate industries and for the three factor endowments (population, physical capital and land area), we have an unbalanced panel of 811 observations per industry on the 45 regions during approximately 1975-95 (Sample A). Second, for the disaggregated manufacturing industries and for the 5 factor endowments (low education, medium education, high education, physical capital and land area), we have an unbalanced panel of 696 observations per industry from approximately 1980 onwards (Sample B). Full details of the composition of each sample are given in Appendix A.

Table 1 presents information on the share of the three aggregate industries in each region's GDP in 1975, 1985 and 1995. We find substantial variation in patterns of production across regions at any one point in time, even at the level of the three aggregate industries. For example, the share of Agriculture in GDP in 1985 varies from 0.03% in Be1 (Brussels) to 11.86% in Esp4 (Centre), while the GDP share of Services in the same year ranges from 81.61% in Be1 (Brussels) to 49.57% in Esp2 (North East). There are also marked changes in patterns of specialization over time. Thus, the GDP share of Agriculture in Esp4 (Centre) falls from 14.72% in 1980 to 5.39% in 1995, while the GDP

[17] We also experiment with using data on arable land area to control for variation in land quality.

[18] See Appendix A for further information concerning the data used.

share of Services in Fra3 (Nord-Pas-de-Calais) rises from 46.68% in 1975 to 67.00% in 1995.

The share of the disaggregated manufacturing industries in GDP also varies substantially, both across regions at a point in time and within regions over time.[19] For example, the GDP share of Metal Products and Machinery (Machine) in Fra7 (Centre-East) in 1985 is almost 3 times larger than that in Fra8 (Mediterranean) and almost 6 times larger than that in Esp6 (South). The GDP share of Chemicals in Esp6 (South) falls by 45% between 1980 and 1994, while the share of Paper in Fra3 (Nord-Pas-de-Calais) rises by 24% over the same period.

Table 2 examines variation in the three broad factor endowments (Population, Capital and Land) across regions at a point in time and within regions over time. The sample includes both UK5 (with a population of more than 16 million in 1985) and Luxembourg (with a population of just over 350,000 in 1985). Land area varies from around 16,000 hectares in Be1 (Brussels) to 21 million hectares in Es4 (Centre). While population declined in some regions, such as UK1 (North), it rose in others, such as UK6 (South-West). All regions exhibit an increase in the real stock of physical capital over time, although the rate of increase varies across regions.

In Table 3 we report regional educational attainment as a percentage of the population for the years 1985 and 1995. It is well known from the labour market literature that the sample period was one of rising educational attainment in European countries (see, for example, Nickell and Bell 1996 and Machin and Van Reenen 1998). With the exception of Be1 (Brussels), all regions in Table 4 experience a rise in the share of the population with high education. However the rate of increase varies substantially, even across regions even within a country. For example, in Esp2 (North-East) the high education share rises by over 70%, while in the neighbouring region of Esp1 (North-West) the proportional rate of increase is approximately 40%. Multiplying the percentage shares in Table 3 by the population levels reported in Table 2, we obtain regions' endowments of low, medium and high education individuals.

[19] In the interests of brevity, the GDP shares for the disaggregated manufacturing industries are not reported in full here. For further details, see the earlier working paper version of this paper (Redding and Vera-Martin 2001).

4 Econometric Specification

Our null and alternative specifications are derived directly from the structure of the theoretical model, as explained above, and are reproduced below:

$$
\begin{aligned}
\textbf{(NULL)} \qquad s_{zjt} &= \sum_t \phi_{jt} d_t + \sum_i \gamma_{ji} \ln v_{zit} + \varepsilon_{zjt} \\
\textbf{(ALT1)} \qquad s_{zjt} &= \eta_{cj} + \sum_t \zeta_{jt} d_t + \sum_i \gamma_{ji} \ln v_{zit} + \omega_{zjt} \\
\textbf{(ALT2)} \qquad s_{zjt} &= \sum_c \sum_t \zeta_{cjt} d_{ct} + \sum_i \gamma_{ji} \ln v_{zit} + \omega_{zjt} \\
\textbf{(ALT3)} \qquad s_{zjt} &= \eta_{zj} + \sum_c \sum_t \zeta_{cjt} d_{ct} + \sum_i \gamma_{ji} \ln v_{zit} + \omega_{zjt}
\end{aligned}
$$

Since all coefficients vary across industries, specifications are estimated separately for each industry, pooling observations across regions and over time.

As we move from the null to each of the alternative specifications, we change the source of variation in the data used to identify the coefficients on factor endowments (γ_{ji}). For example, in **(NULL)**, the inclusion of time dummies in each industry regression means that we abstract from any common trend in factor endowments over time across all regions, and the γ_{ji} are identified from variation across regions at a point in time and differential variation within regions over time. In **(ALT2)**, the inclusion of country-time dummies means that we abstract from any common trend in factor endowments across all regions within a country, and the γ_{ji} are identified from variation across regions within a country at a point in time and differential variation within regions over time. If the assumptions of the HO model are satisfied, the additional terms included in specifications **(ALT1)-(ALT3)** should be statistically insignificant, and the estimated γ_{ji} should remain unchanged, as we move to progressively more general models of relative prices and technology and exploit different sources of variation in the data.

However, as we move from **(ALT2)** to **(ALT3)**, the within groups transformation due to the inclusion of the regional fixed effect (η_{zj}) can greatly exacerbate any attenuation bias from measurement error in the independent variables (see, in particular, Griliches and Hausman 1986). Intuitively, the extent of 'within' or time-series variation in factor endowments due to true variation in the independent variables may be small relative to the variation due to measurement error. This is likely to be a particular problem in the present application because the extent of time-series variation in some of our factor endowments (in particular land area and, to a lesser extent, population) is limited.

We address this problem in two ways. First, we exploit disaggregated data on the educational attainment of the population and on arable land area. The resulting measures of factor endowments control for variation in levels of skills and land quality, and exhibit greater differential variation over time within regions. Second, following Griliches and Hausman (1986), we consider the use of first-differenced estimators. The longer the interval of time over which we difference the data, the greater the amount of true variation in factor endowments relative to that due to measurement error. Hence, the attenuation bias due to measurement error should be smaller using longer differences, and we analyze the results of 10-year difference estimators. We thus obtain a fourth alternative specification:

$$\textbf{(ALT4)} \qquad \triangle_{10} s_{zjt} = \sum_t \zeta_{jt} d_t + \sum_i \gamma_{ji} \triangle_{10} \ln v_{zit} + \psi_{zjt}$$

where differencing eliminates the regional fixed effect (η_{zj}). In taking long differences, we substantially reduce the sample size and, therefore, we concentrate on a specification with only industry-specific coefficients on the time dummies.

In comparing the results of estimating the null and alternative specifications, we also make use of two model specification tests. The first of these focuses on the time-series properties of the model. By construction, the share of sector j in GDP (s_{zjt}) is bounded between 0 and 100 per cent, and is therefore I(0). However, in any finite sample, GDP shares may be I(1). This is particularly true of our sample period (1975-95) which, in general, is characterized by a secular decline in the GDP shares of Agriculture and Manufacturing combined with a secular rise in the share of Services. Similarly, a region's population and physical capital endowments may be I(1). In this case, the static levels regressions **(NULL)-(ALT3)** should be interpreted as cointegrating relationships between a sector's share of GDP and factor endowments. Under this interpretation, the residuals should be I(0) if the assumptions underlying a particular specification are satisfied. Therefore, we make use of the panel data unit root test of Maddala and Wu (1999) to test for the stationarity of the residuals.[20]

Second, neoclassical trade theory assumes that the production technology is constant returns to scale and hence that the revenue function in equation (1) is homogeneous of degree 1 in factor endowments. Therefore, a test of the null hypothesis that the sum of the estimated coefficients on factor endowments is equal to zero provides another model specification test ($\sum_i \gamma_{ji} = 0$).

5 Empirical Results

We begin in Column (1) of Tables 4A-C by reporting the results of estimating the HO specification **(NULL)** for the aggregate industries (Agriculture, Manufacturing and Services) using our 3 broad

[20] The Maddala and Wu or Fisher test statistic is based on the sum of the p-values from conventional Augmented Dickey Fuller (ADF) tests on the residuals for each cross-section unit $z \in Z$. It can be shown that $-2\sum_z \ln P_z$ has a χ^2 distribution with $2Z$ degrees of freedom. This test statistic has a direct intuitive interpretation, is valid for unbalanced panels and has attractive small sample properties (Maddala and Wu 1999). Other analyses of unit roots and cointegration in a panel data context include Im *et al.* (1997), Levin and Lin (1992), Pedroni (1999), Pesaran *et al.* (1998) and Quah (1994).

measures of factor endowments (population, physical capital and land area). We find a statistically significant relationship between regional patterns of production and factor endowments and, from the regression R^2, the HO specification explains some 30-45% of the variation in production patterns across European regions.[21] We also find statistically significant effects of the year dummies, consistent with an important role for common changes in relative prices and technology.

In Column (2) of Tables 4A-C, we relax the assumption of identical relative prices and technology in the first alternative specification **(ALT1)**. The country fixed effects are highly statistically significant, as shown in the first of the F-statistics reported in Column (2). The R^2 of the regression rises substantially, and by more than one third in both Agriculture and Manufacturing. The pattern of estimated coefficients on factor endowments also changes and moves more in line with economic priors concerning factor intensity. For example, in Agriculture the coefficient on physical capital switches from being positive and statistically significant to negative and statistically significant, while in Manufacturing the coefficients on physical capital and population increase by an order of magnitude. The statistical significance of the country fixed effects rejects the null hypothesis of the HO model; the rise in R^2 suggests that the additional terms capturing variation in relative prices and technology are quantitatively important; and the change in the estimated pattern of coefficients suggests the importance of including these controls in identifying the relationship between production patterns and factor endowments.

In Column (3) of Tables 4A-C, we generalize the model of relative prices and technology further and consider specification **(ALT2)** which allows for different cross-country trends in relative prices and technology over time. In the HO specification **(NULL)**, each industry regression includes a set of time dummies. In moving from **(NULL)** to **(ALT2)**, we retain these time dummies (which will capture effects for the omitted country) and augment the regression specification with country-time dummies for all countries except one. The first of the F-statistics in this column of the tables reports the results of a test whether the coefficients on the additional country-year dummies are statistically significantly different from zero. In all industries, the country-year dummies are highly statistically significant, and the HO specification **(NULL)** is again rejected against the more general alternative.

Comparing Columns (2) and (3), the pattern of estimated coefficients on factor endowments remains stable between specifications **(ALT1)** and **(ALT2)**. This suggests that controlling for different cross-country trends in relative prices and technology does not substantially alter the relationship between factor endowments and regional production patterns, and that it is far more important to control for permanent cross-country differences in relative prices and technology (the move from

[21]Except where otherwise indicated, statements about statistical significance refer to the 5% level.

(NULL) to **(ALT1)**).

The values of the estimated coefficients on factor endowments in **(ALT2)** are generally consistent with economic priors. Population endowments are positively correlated with specialization in Services and negatively correlated with specialization in Manufacturing. Greater endowments of physical capital are associated with a higher share of Manufacturing in GDP and a lower share of Agriculture and Services. Land area is positively related to specialization in Agriculture and Manufacturing and negatively related to specialization in Services. With the exception of the coefficient on population in the regression for Agriculture, all estimated coefficients are statistically significant at the 5% level.

Column (4) of Tables 4A-C reports the results of extending the model of relative prices and technology further to include a region-industry fixed effect in specification **(ALT3)**. Here, the pattern of estimated coefficients changes substantially and no longer has a plausible economic interpretation. For example, land area is negatively correlated with the share of Agriculture in GDP, while endowments of physical capital are positively and statistically significantly correlated with specialization in Agriculture. Since there is almost no time-series variation in land area (see Table 3), it is unclear how appropriate or meaningful this econometric specification is. The parameters of interest are being identified from deviations from time means for individual regions, which in all cases are extremely small and in many cases are literally zero. It is plausible that the change in the estimated coefficients between **(ALT2)** and **(ALT3)** is largely driven by measurement error (Griliches and Hausman 1986). We investigate this possibility further below, where we disaggregate factor endowments (thereby introducing more time-series variation) and explore the results of long differences estimation.

Tables 4A-C also report the sum of the estimated coefficients on factor endowments in each industry and the results of a test whether the revenue function is linearly homogenous of degree 1 in factor endowments (a test of the null hypothesis that $\sum_i \gamma_{ji} = 0$). Although the sum of the estimated coefficients is close to zero (in several cases, the order of magnitude is 10^{-2}), the null hypothesis is frequently rejected at conventional levels of statistical significance. There is some evidence of increasing returns to scale in Manufacturing, where the sum of the estimated coefficients is strictly greater than zero in all specifications.

Our other model specification test examines the stationarity of the residuals using the unit root tests of Maddala and Wu (1999). In Agriculture we are able to reject the null hypothesis of a unit root in the residuals in all specifications, while in Services and Manufacturing we are unable to reject the null hypothesis in half the specifications. Taken together, these results provide some evidence of model mis-specification. Two possible explanations for the non-stationarity of the residuals are the omission of information on relevant factor endowments or time-varying regional price and technology differences that have not been controlled for (both of which will be included in the error term).

Table 5 investigates the first of these possibilities by introducing information on educational attainment and land quality. The availability of the educational attainment data reduces the sample size to 696 observations per industry (Sample B).[22] In the interests of brevity, we only report the results for specifications **(ALT2)** and **(ALT3)**. In both cases, as shown in the first F-statistic reported in the table, we reject the HO null against the more general neoclassical alternative, and we begin by considering the estimation results for specification **(ALT2)**.

The estimated coefficients on physical capital are very similar to before, while the arable land coefficients closely resemble those on total land area. In addition, we find highly statistically significant effects of education endowments. Greater endowments of low-education labour are positively and statistically significantly correlated with specialization in Agriculture, while endowments of medium-education labour are negatively and statistically significantly correlated with the share of Agriculture in GDP. There is a positive and statistically significant relationship between endowments of medium-education labour and Manufacturing's share of GDP, while the relationship with endowments of high-education labour is negative and statistically significant. Endowments of high-education labour are positively and statistically significantly linked with specialization in Services, while endowments of medium-education labour are negatively and statistically significantly linked with the share of Services in GDP. This pattern of results is consistent with the idea that Services is skilled-labour intensive relative to Agriculture and Manufacturing.

The introduction of more disaggregated measures of factor endowments increases the regression R^2, which in Manufacturing rises from 0.42 in Column (3) of Table 4B to 0.50 in Column (2) of Table 5. Furthermore, we are now able to reject the null hypothesis of a unit root in the residuals at the 5% level in all three industries. This is consistent with the idea that the non-stationarity of the residuals in the specification with population, physical capital and land area was due to the omission of information on relevant factor endowments. The sum of the estimated coefficients on factor endowments in all three industries is again close to zero, although the null hypothesis that the revenue function is linearly homogenous of degree 1 is typically rejected at the 5% level. The sum of the estimated coefficients in Manufacturing remains strictly greater than zero, again providing some evidence of increasing returns to scale.

The introduction of the region-industry fixed effects in **(ALT3)** again leads to a change in the estimated pattern of coefficients which often no longer have a plausible economic interpretation. For

[22] The model in Tables 4A-C was re-estimated for the reduced sample; this yields very similar results to those reported in the paper.

example, increases in arable land area are negatively (though not statistically significantly) related to specialization in Agriculture. Again, it is plausible that these results are driven by measurement error - the extent of true time-series variation in factor endowments within regions still remains small relative to that due to measurement error, and this is particularly the case for arable land area.[23]

Table 6 investigates this possibility further using the results of long differences estimation over a 10-year time period **(ALT4)**. The long differences estimator enables us to control for unobserved heterogeneity at the regional level, while reducing the magnitude of any attenuation bias induced by measurement error. The pattern of estimated coefficients in Table 6 is similar to that reported for **(ALT2)** in Table 5. For example, arable land area is positively and statistically significantly correlated with the share of Agriculture in GDP and negatively and statistically significantly correlated with the share of Services. The main exception is for the low education endowment where one of the estimated coefficients changes sign.

The constancy of the estimated parameters as one moves from **(ALT1)** to **(ALT2)** in Tables 4A-C, the fact that **(ALT2)** is explicitly concerned with variation across regions within countries, and the support provided by the long differences estimation lead us to select **(ALT2)** as our preferred specification. Throughout the remainder on the paper, we concentrate on the results using the more disaggregated data on factor endowments that control for educational attainment and land quality.

The analysis so far has established that the HO model's assumptions of identical relative prices and technology are rejected against more general alternatives consistent with the neoclassical model of trade. The additional terms capturing regional variation in relative prices and technology are not only highly statistically significant but also quantitatively important. In Manufacturing the regression R^2 rises from 0.29 in **(NULL)** to 0.42 in our preferred alternative specification **(ALT2)**, while in Services the R^2 rises from 0.45 to 0.55. Taken together, the results provide evidence that HO is an incomplete model of patterns of production across European regions.

Nonetheless, factor endowments remain highly statistically significant within each of the alternative specifications, and Table 7 examines their quantitative importance in explaining patterns of production within our preferred alternative specification. The first row of each panel of the table reports the average share of a sector in GDP for the whole sample and for individual countries. The remaining rows of each panel report within sample average absolute prediction errors. These are defined in proportional terms as $|s_{zjt} - \hat{s}_{zjt}| / s_{zjt}$, where a hat above a variable indicates a predicted value.[24] The prediction errors reported in the table differ in terms of how predicted shares of GDP

[23] The time-series variation in arable land area, though larger than that in total land area, remains small.

[24] The model's predictions for output levels can be obtained by multiplying predicted GDP shares by actual GDP. Proportional prediction errors for output are therefore exactly the same as for shares of sectors in GDP (one is multiplying both the numerator and denominator of the formula in the text by actual GDP).

are calculated. The first and simplest measure uses the fitted values from **(ALT2)**, and provides an overall indication of the model's within-sample predictive ability. The second measure takes the parameter estimates from **(ALT2)** but only uses the country-year dummies to predict shares of sectors in GDP. Comparing the second and first measures reveals the contribution of factor endowments to reducing the model's within-sample prediction error. The third measure takes the parameter estimates from **(ALT2)** but only uses the 5 factor endowments to predict shares of sectors in GDP. Comparing the third and first measures reveals the contribution of the terms capturing relative prices and technology to reducing the model's within-sample prediction error.

The model's overall average prediction error across regions and years in Manufacturing is 13%, which compares favorably with the average prediction error across disaggregated manufacturing industries in Harrigan (1995) using country-level data (38%) and the average prediction errors reported using regional data in Bernstein and Weinstein (2002). For individual countries, the average prediction error within Manufacturing varies from 6% in Belgium to 18% in the Netherlands. Looking across industries, the model is most successful at explaining European regional production patterns in Services and Manufacturing. For the whole sample and all countries except the Netherlands, we find the same ranking of industries in terms of (increasing) average prediction errors: from Services, through Manufacturing, to Agriculture.

Factor endowments remain quantitatively important in explaining variation in production patterns across European regions. The model's average prediction error rises by a factor of more than 3 in Manufacturing and more than doubles in Services if information on factor endowments is excluded. In Agriculture and Manufacturing, excluding the country-year dummies has a roughly similar effect to excluding factor endowments, suggesting that these two sets of considerations make roughly equal contributions towards explaining variation in specialization patterns. In Services, the country-year dummies are much more important. In general equilibrium, variation in relative prices affects the share of all sectors in GDP (the country-year dummies are important in all sectors), but the finding of the largest effects in Services is consistent with this sector being the least tradeable.

One of the features that makes our sample period interesting is that it is characterized by increasing European integration. In Table 8, we examine the model's average absolute prediction errors over time. Has the process of closer integration weakened the relationship between regions' patterns of production and their factor endowments, so that we observe an increase in average prediction errors over time? Since the country-year dummies in **(ALT2)** control for any country-specific changes in patterns of production over time, we are explicitly concerned here with how increasing integration has affected the relationship between production and endowments across regions *within countries.* From

Table 8, we find no systematic increase or decrease in average prediction errors over time. Across all regions and years, the average prediction error falls in Services and remains broadly constant in Manufacturing and Agriculture.

Finally, it is frequently asserted that factor endowments explain specialization and trade at the aggregate level in industries such as Agriculture, Manufacturing and Services, while other considerations, including imperfect competition and increasing returns to scale, are more important for specialization and trade within these aggregate industries. This hypothesis is implicit in the construction of theoretical models of inter and intra-industry trade, such as Helpman and Krugman (1985). The same assumption is made in empirical work by Davis and Weinstein (1999), (2002). The present dataset and empirical framework may be used to shed light on whether this hypothesis holds for European regions. In addition to the aggregate industries considered above, we also estimate the model for disaggregated industries within the manufacturing sector, and the results are reported in Tables C1A and C1B of Appendix C.

Factor endowments are again found to play a statistically significant role in explaining production structure in European regions. For example, physical capital is positively and statistically significantly related to the share of Chemicals, Machinery and Transport Equipment in a region's GDP. Medium education is positively and statistically significantly correlated with specialization in Metals, Machinery and Transport Equipment. However, in all industries, the HO specification **(NULL)** is rejected against the more general neoclassical alternative **(ALT2)** at the 5% level of statistical significance

Table 9 examines average absolute prediction errors at the disaggregated level. In 10 of the 11 manufacturing industries and for every 5-year period considered, average prediction errors across regions and time are higher than those reported for manufacturing as a whole in Table 8 (the exception is the Construction industry). Considering all 11 disaggregated industries together, the average prediction error across regions during 1985-90 was 47%, compared with an average prediction error across the 3 aggregated industries in Table 8 of 31% during the same period. This provides evidence that factor endowments are indeed more successful at explaining patterns production at the aggregate level (Agriculture, Manufacturing and Services) than in disaggregated industries within the manufacturing sector.

These findings concerning the model's predictive ability are consistent with our theoretical approach. At the aggregate level, there are at least as many factors of production as the number of goods ($M \geq N$), which we noted earlier is a condition for the revenue function to be twice continuously differentiable. Whereas for the 11 disaggregated manufacturing industries, there are more goods than factors of production ($N > M$). One theory-consistent explanation for the larger disaggregated prediction errors is, therefore, that there is a degree of indeterminacy in the production of individual manufacturing industries at the regional level. Another theory-consistent explanation is that there are larger price and technology differences across regions within individual disaggregated manufacturing industries that are not being captured in the right-hand side variables.

At the disaggregated level, we also find no systematic trend in the average absolute prediction errors over time, so that there is again no evidence that increasing European integration has weakened the relationship between factor endowments and production across regions within countries.

6 Conclusions

This paper has analyzed the relationship between production patterns and factor endowments using data on a panel of 14 industries in 45 regions from 7 European countries since 1975. Under the assumptions of the Heckscher-Ohlin (HO) model of identical relative prices and technology, we derived a general equilibrium structural relationship between the share of a sector in GDP and factor endowments. The HO model is a special case of the more general neoclassical model of trade, which allows for regional variation in relative prices and technology. We compared the empirical performance of the HO null against a series of alternative specifications derived from the neoclassical model and including progressively more general models of relative prices and technology.

The use of European regional data enables us to abstract from many of the considerations that have been proposed as explanations for the disappointing empirical performance of HO theory at the country-level. For example, both measurement error and technology differences are likely to be much smaller across regions within Europe than for a cross-section of developed and developing countries. If factor endowments are mobile across regions, the general equilibrium relationships that we estimate have both a demand-side and a supply-side interpretation. Irrespective of whether the relationships are demand-side, supply-side or a combination of both, we are able to test between the null and alternative specifications, and we are able to examine the respective contributions of factor endowments and other considerations to explaining patterns of production.

For both aggregate industries (Agriculture, Manufacturing and Services) and disaggregated manufacturing industries, the HO null is rejected against more general neoclassical alternatives that allow for regional variation in relative prices and technology. Nevertheless, within each of the alternative specifications considered, factor endowments remain highly statistically significant and make an important contribution to explaining patterns of production. Excluding information on factor endowments in our preferred alternative specification increases within-sample prediction errors for

Manufacturing by a factor of more than 3. The pattern of estimated coefficients across industries accords with economic priors. For example, endowments of physical capital are positively correlated with the share of Manufacturing in GDP and negatively correlated with the share of Agriculture and Services.

Factor endowments are more successful at explaining production structure at the aggregate level (Agriculture, Manufacturing and Services) than in disaggregated industries within manufacturing, a finding that is consistent with the predictions of theory. The large number of disaggregated manufacturing industries relative to the number of factor endowments suggests the possibility of production indeterminacy, and regional differences in relative prices and technology may be particularly large in individual manufacturing industries. At both the aggregate and disaggregate level, we find no evidence that the process of increasing European integration has weakened the relationship between factor endowments and production across regions within countries.

References

Amiti, M (1999) 'Specialization Patterns in Europe', *Weltwirtschaftliches Archiv*, 135(4), 573-593.

Barro, R and Sala-i-Martin, X (1995) *Economic Growth*, McGraw-Hill.

Bernard, A and Jones, C (1996) 'Comparing Apples to Oranges: Productivity Convergence and Measurement across Industries and Countries', *American Economic Review*, December, 1216-38.

Bernstein, J and Weinstein, D (2002) 'Do Endowments Determine the Location of Production? Evidence from National and International Data', *Journal of International Economics*, 56(1), 55-76.

Bowen, H, Leamer, E, and Sveikauskas, L (1987) 'Multicountry, Multifactor Tests of the Factor Abundance Theory', *American Economic Review*, 77, 791-809.

Brulhart, M (2000) 'Evolving Geographical Specialization of European Manufacturing Industries', mimeograph, University of Lausanne.

Cameron, G and Muellbauer, J (1998) 'The Housing Market and Regional Commuting and Migration Costs', *Scottish Journal of Political Economy*, 45(4), 420-46.

Davis, D, Weinstein, D, Bradford, S, and Shimpo, K (1997) 'Using International and Japanese Regional Data to Determine When the Factor Abundance Theory of Trade Works', *American Economic Review*, 87(3), 421-46

Davis, D and Weinstein, D (1999) 'Economic Geography and Regional Production Structure: An Empirical Investigation', *European Economic Review*, 43, 379-407.

Davis, D and Weinstein, D (2001) 'An Account of Global Factor Trade', *American Economic Review*, 91(5), 1423-53.

Davis, D and Weinstein, D (2002) 'Market Access, Economic Geography, and Comparative Advantage: An Empirical Assessment', *Journal of International Economics*, forthcoming.

Dixit, A and Norman, V (1980) *The Theory of International Trade*, Cambridge: Cambridge University Press.

Ethier, W (1984) 'Higher Dimensional Issues in Trade Theory', Chapter 3 in (eds) Jones, R and Kenen, P, *Handbook of International Economics*, vol I, 131-84.

Gabaix, X (1997) 'The Factor Content of Trade: A Rejection of the Heckscher-Ohlin-Leontief Hypothesis', Harvard University, mimeo.

Gandal, N, Hanson, G, and Slaughter, M (2002) 'Technology, Trade, and Adjustment to Immigration in Israel', *European Economic Review*, forthcoming.

Griffith, R, Redding, S, and Van Reenen, J (2000) 'Mapping the Two Faces of R&D: Productivity Growth in a Panel of OECD Manufacturing Industries', *CEPR Discussion Paper*, 2457.

Griliches, Z and Hausman, J (1986) 'Errors in Variables in Panel Data', *Journal of Econometrics*, 31, 93-118.

Hanson, H and Slaughter, M (2002) 'Labor Market Adjustment in Open Economies: Evidence from US States', *Journal of International Economics*, forthcoming.

Harrigan, J (1995) 'Factor Endowments and the International Location of Production: Econometric Evidence for the OECD, 1970-85', *Journal of International Economics*, 39, 123-41.

Harrigan, J (1997) 'Technology, Factor Supplies, and International Specialization: Estimating the Neoclassical Model', *American Economic Review*, 87(4), 475-94.

Harrigan, J and Zakrajsek, E (2000) 'Factor Supplies and Specialization in the World Economy', *NBER Working Paper*, 7848.

Helpman, E (1984) 'Increasing Returns, Imperfect Markets, and Trade Theory', Chapter 7 in (eds) Jones, R and Kenen, P, *Handbook of International Economics*, vol I, 325-65.

Helpman, E and Krugman, P (1985) *Market Structure and Foreign Trade*, MIT Press.

Im, K, Pesaran, M, and Shin, Y (2002) 'Testing for Unit Roots in Heterogeneous Panels', *Journal of Econometrics*, forthcoming.

Kohli, U (1991) *Technology, Duality and Foreign Trade*, Ann Arbor: University of Michigan Press.

Krugman, P (1991) *Geography and Trade*, MIT Press.

Leamer, E (1984) *Sources of Comparative Advantage: Theories and Evidence*, MIT Press.

Levin, A and Lin, C (1992) 'Unit Root Tests in Panel Data: Asymptotic and Finite-sample Properties', *University of California, San Diego, Discussion Paper*, 92-23.

Machin, S and Van Reenen, J (1998) 'Technology and Changes in Skill Structure: Evidence from Seven OECD Countries', *Quarterly Journal of Economics*, November, 1215-44.

Maddala, G and Wu, S (1999) 'A Comparative Study of Unit Root Tests with Panel Data and a New Simple Test', *Oxford Bulletin of Economics and Statistics*, Special Issue, 631-52.

McCormick, B (1997) 'Regional Unemployment and Labour Mobility in the United Kingdom', *European Economic Review*, 41(3-5), 581-89.

Midelfart-Knarvik, K, Overman, H, Redding, S and Venables, A (2000) 'The Location of European Industry', *Economic Papers no. 142*, European Commission, D-G for Economic and Financial Affairs, Brussels.

Midelfart-Knarvik, K, Overman, H, and Venables, A (2000) 'Comparative Advantage and Economic Geography', *CEPR Discussion Paper*, 2618.

Molle, W. (1996) 'The Regional Economic Structure of the European Union: an Analysis of Long-Term Developments',. in (ed) Peschel, K *Regional Growth and Regional Policy Within the Framework of European Integration*, Physica-Verlag.

Nickell, S and Bell, B (1996) 'Changes in the Distribution of Wages and Unemployment in OECD Countries', *American Economic Review*, 86(2), 302-8.

Nickell, S, Redding, S, and Swaffield, J (2001) 'Educational Attainment, Labour Market Institutions, and the Structure of Production', *CEPR Discussion Paper*, 3068.

Pedroni, P (1999) 'Critical Values for Cointegration Tests in Heterogeneous Panels with Multiple Regressors', *Oxford Bulletin of Economics and Statistics*, Special Issue, 653-70.

Pesaran, M, Shin, Y, and Smith, R (1999) 'Pooled Mean Group Estimation of Dynamic Heterogeneous Panels', *Journal of the American Statistical Association*, 94, 621-34.

Overman, H, Redding, S, and Venables, A (2002) 'The Economic Geography of Trade Production and Income: a Survey of Empirics', forthcoming in (eds) J Harrigan and E Kwon-Choi, *Handbook of International Trade*, Basil Blackwell.

Proudman, J and Redding, S (1998) 'Persistence and Mobility in International Trade', Chapter 2 in (eds) Proudman, J and Redding, S, *Openness and Growth*, Bank of England.

Proudman, J and Redding, S (2000) 'Evolving Patterns of International Trade', *Review of International Economics*, August, 8(3), pages 373-96.

Quah, D (1994) 'Exploiting Cross-section Variation for Unit Root Inference in Dynamic Data', *Economics Letters*, 44, 9-19.

Redding, S (2002) 'Specialization Dynamics', *Journal of International Economics,* 58(2), 299-334.

Redding, S and Vera-Martin, M (2001) 'Factor Endowments and Production in European Regions', *CEP Discussion Paper*, 501, London School of Economics.

Schott, P (2002) 'One Size Fits All? Heckscher-Ohlin Specialization in Global Production', *American Economic Review*, Forthcoming.

Trefler, D (1995) 'The Case of the Missing Trade and Other Mysteries', *American Economic Review*, 85(5), pages 1029-46.

Vera-Martin, M (2000) 'Specialization in European Regions', London School of Economics, mimeograph.

Table 1: Shares of Agriculture, Manufacturing and Services in GDP (%)

Region	Year	Agric	Manu	Serv
Belgium				
Be1	1975	0.01	28.58	71.42
	1985	0.03	18.35	81.61
	1995	0.02	15.75	84.23
Be2	1975	3.64	42.45	53.92
	1985	2.64	37.39	59.96
	1995	1.52	34.27	64.21
Be3	1975	3.91	39.93	56.16
	1985	3.00	32.24	64.76
	1995	1.76	27.36	70.88
Spain				
Esp1	1980	9.35	39.22	51.43
	1985	8.07	39.29	52.64
	1995	4.82	34.45	60.73
Esp2	1980	5.93	48.12	45.95
	1985	4.59	45.84	49.57
	1995	2.25	42.12	55.63
Esp3	1980	0.55	30.58	68.86
	1985	0.32	28.61	71.08
	1995	0.17	25.26	74.57
Esp4	1980	14.72	34.61	50.67
	1985	11.86	36.10	52.04
	1995	5.39	34.50	60.11
Esp5	1980	4.24	42.19	53.56
	1985	3.00	39.44	57.56
	1995	1.57	34.55	63.87
Esp6	1980	10.91	33.13	55.97
	1985	10.84	29.25	59.91
	1995	6.15	27.91	65.95
Esp7	1980	8.25	21.38	70.37
	1985	4.80	18.41	76.79
	1995	2.06	18.56	79.38
France				
Fra1	1975	0.68	34.87	64.45
	1985	0.40	29.50	70.10
	1995	0.18	22.54	77.28
Fra2	1975	8.73	43.58	47.69
	1985	7.41	36.07	56.52
	1995	4.27	32.69	63.04
Fra3	1975	4.04	49.28	46.68
	1985	2.62	35.80	61.58
	1995	1.35	31.66	67.00
Fra4	1975	4.45	47.50	48.05
	1985	3.71	37.48	58.81
	1995	2.60	34.66	62.74
Fra5	1975	11.38	37.48	51.14
	1985	7.98	29.24	62.78
	1995	5.24	27.08	67.69
Fra6	1975	8.57	37.15	54.28
	1985	6.67	30.56	62.77
	1995	4.41	24.37	71.22
Fra7	1975	4.30	46.46	49.24
	1985	3.11	36.85	60.04
	1995	2.43	32.17	65.40
Fra8	1975	5.98	32.96	61.06
	1985	4.27	25.21	70.52
	1995	3.24	20.26	76.51

Region	Year	Agric	Manu	Serv
Italy				
Ita1	1975	4.46	48.20	47.34
	1985	3.12	41.17	55.71
	1995	2.42	36.87	60.71
Ita2	1975	2.87	55.18	41.95
	1985	2.05	45.49	52.45
	1995	1.56	41.21	57.23
Ita3	1975	6.03	45.82	48.15
	1985	4.38	40.96	54.66
	1995	3.13	36.71	60.17
Ita4	1975	8.72	47.13	44.15
	1985	5.75	40.87	53.38
	1995	3.70	36.94	59.36
Ita5	1975	5.32	45.34	49.34
	1985	3.36	41.09	55.55
	1995	2.64	34.62	62.75
Ita6	1975	4.40	26.02	69.57
	1985	2.59	25.09	72.32
	1995	1.62	21.24	77.14
Ita7	1975	10.70	39.19	50.11
	1985	6.72	34.32	58.95
	1995	4.53	32.04	63.43
Ita8	1975	10.26	31.29	58.45
	1985	5.38	27.71	66.91
	1995	3.33	24.30	72.36
Ita9	1975	14.17	31.02	54.81
	1985	9.42	27.81	62.77
	1995	6.49	24.49	69.02
Itaa	1975	12.56	29.54	57.90
	1985	9.36	28.33	62.32
	1995	5.73	21.47	72.80
Itab	1975	9.09	36.40	54.51
	1985	5.96	33.64	60.39
	1995	4.11	27.37	68.52
Luxembourg				
Lux	1975	3.24	39.18	57.57
	1985	2.36	34.04	63.60
	1995	1.21	31.61	67.19
Netherlands				
Nld1	1975	7.55	36.91	55.54
	1985	4.33	35.08	60.58
	1995	4.45	38.53	57.02
Nld2	1975	7.24	35.21	57.55
	1985	6.12	29.77	64.11
	1995	4.19	27.24	68.58
Nld3	1975	3.73	32.48	63.80
	1985	3.29	28.23	68.48
	1995	2.71	23.64	73.65
Nld4	1975	4.70	43.21	52.09
	1985	5.48	37.25	57.27
	1995	3.55	32.62	63.83

Notes: figures may not sum exactly to 100 due to rounding

Table 1 (cont.): Shares of Agriculture, Manufacturing and Services in GDP (%)

Region	Year	Agric	Manu	Serv
United Kingdom				
UK1	1975	2.04	50.04	47.93
	1980	1.58	47.29	51.12
	1985	1.37	41.37	57.25
UK2	1975	2.31	46.46	51.23
	1980	1.92	44.88	53.21
	1985	1.43	41.40	57.17
UK3	1975	2.76	48.36	48.88
	1980	2.76	46.45	50.79
	1985	1.59	44.29	54.11
UK4	1975	6.56	36.12	57.32
	1980	5.78	35.93	58.29
	1985	3.07	36.02	60.91
UK5	1975	0.82	32.14	67.03
	1980	0.85	32.51	66.65
	1985	0.48	29.46	70.06
UK6	1975	3.27	36.86	59.87
	1980	3.23	35.60	61.17
	1985	2.16	34.36	63.48
UK7	1975	1.51	48.94	49.55
	1980	1.71	46.34	51.94
	1985	1.23	43.64	55.13
UK8	1975	0.87	45.53	53.60
	1980	0.70	45.46	53.84
	1985	0.53	44.32	55.15
UK9	1975	2.81	45.57	51.61
	1980	2.88	45.14	51.98
	1985	2.48	45.93	51.59
UKa	1975	2.98	43.73	53.28
	1980	2.39	42.17	55.44
	1985	1.69	38.82	59.49
UKb	1975	3.40	43.37	53.23
	1980	3.30	37.21	59.49
	1985	2.89	36.00	61.11

Notes: figures may not sum exactly to 100 due to rounding

Table 2: Factor Endowments of European Regions

Region	Year	Pop	Cap	Land
Belgium				
Be1	1975	967.38	6721.54	16.20
	1985	961.10	7945.64	16.10
	1995	944.90	9068.48	16.10
Be2	1975	5400.21	11579.08	1351.10
	1985	5646.70	21646.15	1351.20
	1995	5852.00	37943.63	1351.20
Be3	1975	3160.45	6758.03	1684.80
	1985	3197.80	9619.44	1684.40
	1995	3307.90	14563.69	1684.40
Spain				
Esp1	1975	4210.96	27446.60	4528.80
	1985	4443.00	39729.87	4532.80
	1995	4298.00	54444.34	4536.20
Esp2	1975	3855.28	34242.75	7037.40
	1985	4088.35	44469.33	7038.60
	1995	3993.60	63817.20	7034.30
Esp3	1975	4345.41	31732.97	799.50
	1985	4824.05	39215.81	799.50
	1995	5040.40	65441.33	802.80
Esp4	1975	4947.70	33809.81	21492.30
	1985	5217.08	54417.75	21483.50
	1995	5170.50	74640.40	21483.60
Esp5	1975	9490.14	83441.64	6020.50
	1985	10169.84	108768.10	6013.40
	1995	10594.50	169036.20	6014.80
Esp6	1975	6667.77	38905.89	9858.50
	1985	7449.77	59331.64	9858.70
	1995	8197.30	93269.52	9867.60
Esp7	1975	1229.36	7540.54	746.60
	1985	1389.59	11781.26	750.00
	1995	1521.40	20202.67	748.00
France				
Fra1	1975	9899.95	177317.30	1196.50
	1985	10345.20	272393.20	1196.50
	1995	10703.70	396078.60	1196.50
Fra2	1975	8877.92	147434.40	14659.90
	1985	9452.40	206561.20	14659.90
	1995	9888.50	231949.60	14659.90
Fra3	1975	3854.59	40983.74	1245.10
	1985	3910.90	61957.58	1245.10
	1995	3821.90	73979.10	1245.10
Fra4	1975	4694.08	77341.84	4830.90
	1985	4670.80	108023.80	4830.90
	1995	4858.40	123198.00	4830.90
Fra5	1975	6465.02	83522.45	8585.60
	1985	6927.40	124273.00	8585.60
	1995	7589.10	148911.20	8585.60
Fra6	1975	5014.20	75540.89	10449.00
	1985	5607.20	108094.20	10449.00
	1995	5932.90	120266.00	10449.00
Fra7	1975	5884.79	87949.34	7113.60
	1985	6388.30	131157.80	7113.60
	1995	6765.80	154844.70	7113.60
Fra8	1975	5240.18	57037.51	6828.20
	1985	5627.40	96296.46	6828.20
	1995	6775.30	124957.20	6828.20

Region	Year	Pop	Cap	Land
Italy				
Ita1	1975	6431.26	65155.07	3407.60
	1985	6199.00	101068.00	3407.70
	1995	5978.60	131306.60	3407.90
Ita2	1975	8665.99	94042.59	2385.03
	1985	8752.70	154774.10	2385.70
	1995	8786.70	202913.60	2387.30
Ita3	1975	6229.93	67239.92	3982.47
	1985	6344.90	109107.30	3983.10
	1995	6407.30	147813.50	3982.70
Ita4	1975	3864.12	38671.77	2212.30
	1985	3893.20	65189.58	2212.30
	1995	3866.70	82760.49	2212.30
Ita5	1975	5642.75	53216.01	4114.13
	1985	5750.30	88498.08	4114.20
	1995	5714.00	105956.00	4114.20
Ita6	1975	4823.32	37194.79	1720.30
	1985	5008.70	70660.65	1720.30
	1995	5099.10	115990.20	1720.30
Ita7	1975	1494.95	16654.18	1523.20
	1985	1555.70	26232.84	1523.20
	1995	1579.20	30796.81	1523.20
Ita8	1975	5147.29	31031.16	1359.50
	1985	5557.10	59298.32	1359.50
	1995	5687.10	74722.96	1359.50
Ita9	1975	6255.12	44407.48	4442.00
	1985	6620.90	75853.26	4442.00
	1995	6654.60	91471.69	4442.00
Itaa	1975	4739.18	28564.20	2570.80
	1985	4973.00	54135.88	2570.80
	1995	5000.30	67786.23	2570.90
Itab	1975	1504.68	14526.73	2409.00
	1985	1607.20	24124.79	2409.00
	1995	1639.90	30440.95	2409.00
Luxembourg				
Lux	1975	351.73	5928.26	258.60
	1985	355.90	8309.38	258.60
	1995	402.50	13997.32	256.80
Netherlands				
Nld1	1975	1465.86	18254.55	904.50
	1985	1553.87	29099.54	1070.00
	1995	1593.90	37060.60	1138.80
Nld2	1975	2579.90	26835.04	1021.10
	1985	2877.51	45296.34	1020.10
	1995	3129.40	61913.82	1097.60
Nld3	1975	6351.84	75012.20	1037.80
	1985	6597.38	125395.70	1123.50
	1995	7099.50	167428.80	1187.10
Nld4	1975	2925.58	33294.70	731.40
	1985	3124.39	53328.02	731.50
	1995	3350.20	74382.43	729.10

Notes: Pop is population in thousands of people, Cap is real physical capital stock in 1990 millions of ECUs, Land is land area in thousands of hectares. See Appendix A for further details.

Table 2 (cont.) : Factor Endowments of European regions

Region	Year	Pop	Cap	Land
United Kingdom				
UK1	1975	3125.56	13258.54	1540.03
	1985	3051.70	21612.87	1540.10
	1995	3055.20	27599.95	1542.10
UK2	1975	4876.12	20716.00	1541.80
	1985	4845.40	32875.40	1542.00
	1995	4959.20	41394.54	1542.10
UK3	1975	3728.18	14490.71	1561.00
	1985	3851.70	24247.27	1563.00
	1995	4063.60	34597.66	1563.00
UK4	1975	1763.64	11419.70	1256.57
	1985	1934.10	17298.85	1257.30
	1995	2092.00	19389.30	1257.30
UK5	1975	16688.35	87776.61	2722.27
	1985	16880.90	134179.30	2722.20
	1995	17570.20	155221.40	2722.70
UK6	1975	4162.70	16877.93	2383.00
	1985	4407.40	28854.15	2385.00
	1995	4711.50	37562.04	2385.00
UK7	1975	5133.62	15359.96	1301.30
	1985	5127.50	27263.04	1301.30
	1995	5231.80	39124.63	1301.30
UK8	1975	6498.89	21552.36	731.43
	1985	6305.70	35511.66	733.10
	1995	6323.10	48322.58	734.40
UK9	1975	2764.09	6902.18	2076.60
	1985	2777.30	13993.15	2076.80
	1995	2868.20	22284.25	2076.60
UKa	1975	5122.10	21536.84	7877.13
	1985	5052.30	33158.39	7878.30
	1995	5051.00	43721.04	7878.30
UKb	1975	1519.85	6602.66	1412.07
	1985	1535.10	10267.07	1412.00
	1995	1598.80	13058.63	1412.20

Notes: Pop is population in thousands of people, Cap is real physical capital stock in 1990 millions of ECUs, Land is land area in thousands of hectares. See Appendix A for further details.

Table 3: Educational Attainment by Region (% of total population)

Region	Year	Low	Med	High
Belgium				
Be1	1985	53.35	41.55	5.10
	1995	35.97	59.68	4.35
Be2	1985	50.97	45.35	3.67
	1995	41.90	52.51	5.59
Be3	1985	51.07	45.27	3.66
	1995	37.49	57.20	5.31
Spain				
Esp1	1985	65.31	28.52	6.18
	1995	44.46	46.94	8.60
Esp2	1985	59.83	32.45	7.72
	1995	38.35	48.30	13.34
Esp3	1985	47.22	40.70	12.08
	1995	33.51	50.17	16.32
Esp4	1985	67.42	26.44	6.14
	1995	48.11	42.43	9.46
Esp5	1985	67.42	30.49	6.13
	1995	39.36	50.94	9.71
Esp6	1985	70.07	24.53	5.39
	1995	49.28	43.28	7.45
Esp7	1985	65.39	28.19	6.42
	1995	43.66	46.23	10.11
France				
Fra1	1985	46.13	34.18	19.44
	1994	36.50	40.50	23.00
Fra2	1985	61.95	35.98	9.36
	1994	40.00	38.50	21.50
Fra3	1985	65.91	35.08	7.92
	1994	49.00	38.50	12.50
Fra4	1985	55.36	38.68	10.80
	1994	40.50	44.00	15.50
Fra5	1985	55.36	38.68	10.08
	1994	40.00	45.00	15.00
Fra6	1985	48.77	40.48	12.96
	1994	35.50	46.00	18.50
Fra7	1985	50.09	39.58	13.68
	1994	36.32	44.28	19.40
Fra8	1985	56.68	36.88	12.24
	1994	41.79	41.29	16.92
Italy				
Ita1	1985	73.54	21.43	5.06
	1995	56.55	34.64	8.81
Ita2	1985	74.03	20.84	5.21
	1995	54.46	35.32	10.22
Ita3	1985	77.29	18.46	4.54
	1995	56.92	35.54	7.54
Ita4	1985	71.98	22.00	5.97
	1995	54.89	35.79	9.33
Ita5	1985	73.19	21.36	5.47
	1995	56.84	34.02	9.14

Region	Year	Low	Med	High
Italy (cont.)				
Ita6	1985	60.43	29.38	9.41
	1995	46.34	39.99	13.67
Ita7	1985	71.65	21.99	6.30
	1995	56.26	34.15	9.59
Ita8	1985	73.40	19.98	6.71
	1995	57.98	31.73	10.29
Ita9	1985	73.91	20.45	5.72
	1995	60.67	29.67	9.67
Itaa	1985	70.94	21.73	7.26
	1995	59.83	30.07	10.10
Itab	1985	77.40	17.16	5.79
	1995	65.70	27.35	6.96
Luxembourg				
Lux	1985	51.07	45.27	3.66
	1990	37.49	57.20	5.31
Netherlands				
Nld1	1985	29.75	62.33	9.03
	1995	14.86	68.53	16.79
Nld2	1985	27.76	62.77	9.33
	1995	14.53	68.48	16.95
Nld3	1985	28.33	58.68	12.02
	1995	14.80	62.88	22.12
Nld4	1985	29.06	61.16	9.47
	1995	15.08	67.58	17.08
United Kingdom				
UK1	1985	40.48	44.32	3.18
	1994	25.51	68.37	5.80
UK2	1985	38.39	46.85	3.56
	1994	24.42	68.10	7.14
UK3	1985	34.89	47.15	4.43
	1994	24.81	67.38	7.60
UK4	1985	36.31	48.60	4.66
	1994	21.79	70.63	7.49
UK5	1985	27.98	53.17	7.41
	1994	19.44	68.50	11.71
UK6	1985	41.86	40.71	4.06
	1994	20.90	70.24	8.63
UK7	1985	34.77	47.92	4.14
	1994	27.67	64.79	7.07
UK8	1985	37.25	47.46	4.26
	1994	25.09	66.89	7.77
UK9	1985	43.44	43.44	5.10
	1994	26.15	66.20	7.54
UKa	1985	55.00	39.80	3.81
	1994	20.30	70.60	8.86
UKb	1985	55.62	39.18	3.97
	1994	35.40	56.13	8.05

Notes: Low is no or primary education; Medium is secondary and/or vocational qualifications; High is college degree of equivalent. Figures may not sum exactly to 100 due to rounding. See Appendix A for further details concerning the data used.

Table 4A: Factor Endowments and Specialization in Agriculture

GDP share	(1)	(2)	(3)	(4)
Obs	811	811	811	811
Years	1975-95	1975-95	1975-95	1975-95
Capital	0.004**	-0.022**	-0.019**	0.011**
	(0.0013)	*(0.0035)*	*(0.0038)*	*(0.0040)*
Population	-0.021**	0.002	-0.002	-0.129**
	(0.0018)	*(0.0040)*	*(0.0043)*	*(0.0153)*
Land	0.017**	0.016**	0.016**	-0.055**
	(0.0008)	*(0.0009)*	*(0.0010)*	*(0.0159)*
Sample	A	A	A	A
Specification	**(NULL)**	**(ALT1)**	**(ALT2)**	**(ALT3)**
Year dummies	yes	yes		
Country effects		yes		
Cty-year dummies			yes	yes
Region effects				yes
Prob>F(NULL-ALT)	N/A	0.0000	0.0000	0.0000
Prob>F(ALL)	0.0000	0.0000	0.0000	0.0000
R-squared	0.40	0.63	0.65	0.96
Sum of Coeff.	0.0003	-0.0046	-0.0050	-0.1730
Linear Homog	*(0.8061)*	*(0.0000)*	*(0.0000)*	*(0.0000)*
(p-value)	Accept	Reject	Reject	Reject
Maddala-Wu	*(0.0188)*	*(0.0389)*	*(0.0263)*	*(0.0002)*
(p-value)	Reject	Reject	Reject	Reject

Notes: Prob>F(NULL-ALT) is the p-value for an F-test of the null hypothesis that the coefficients on the variables excluded from specification (NULL) but incuded in the alternative specification are equal to 0. Prob>F(ALL) is the p-value for the conventional F-test that the coefficients on all independent variables are equal to zero. Sum of Coeff. is the sum of the estimated coefficients on factor endowments. Linear Homog. is the p-value for a test of the null hypothesis that the sum of the estimated coefficients on factor endowments is equal to zero. Maddala-Wu is the p-value for the Maddala and Wu (1999) panel data test of the null hypothesis that the residuals have a unit root. Huber-White heteroscedasticity robust standard errors in parentheses. ** denotes significance at the 5% level, * denotes significance at the 10% level.

Table 4B: Factor Endowments and Specialization in Manufacturing

GDP share	(1)	(2)	(3)	(4)
Obs	811	811	811	811
Years	1975-95	1975-95	1975-95	1975-95
Capital	-0.005	0.071**	0.073**	0.043**
	(0.0040)	*(0.0116)*	*(0.0133)*	*(0.0082)*
Population	-0.008	-0.079**	-0.082**	0.205**
	(0.0058)	*(0.0124)*	*(0.0140)*	*(0.0322)*
Land	0.021**	0.032**	0.033**	-0.130*
	(0.0023)	*(0.0027)*	*(0.0029)*	*(0.0739)*
Sample	A	A	A	A
Specification	**(NULL)**	**(ALT1)**	**(ALT2)**	**(ALT3)**
Year dummies	yes	yes		
Country effects		yes		
Cty-year dummies			yes	yes
Region effects				yes
Prob>F(NULL-ALT)	N/A	0.0000	0.0000	0.0000
Prob>F(ALL)	0.0000	0.0000	0.0000	0.0000
R-squared	0.29	0.40	0.42	0.97
Sum of Coeff.	0.0082	0.0239	0.0240	0.1180
Linear Homog	*(0.0133)*	*(0.0000)*	*(0.0000)*	*(0.1625)*
(p-value)	Reject	Reject	Reject	Accept
Maddala-Wu	*(0.0040)*	*(0.0020)*	*(0.1949)*	*(0.1779)*
(p-value)	Reject	Reject	Accept	Accept

Notes: Prob>F(NULL-ALT) is the p-value for an F-test of the null hypothesis that the coefficients on the variables excluded from specification (NULL) but incuded in the alternative specification are equal to 0. Prob>F(ALL) is the p-value for the conventional F-test that the coefficients on all independent variables are equal to zero. Sum of Coeff. is the sum of the estimated coefficients on factor endowments. Linear Homog. is the p-value for a test of the null hypothesis that the sum of the estimated coefficients on factor endowments is equal to zero. Maddala-Wu is the p-value for the Maddala and Wu (1999) panel data test of the null hypothesis that the residuals have a unit root. Huber-White heteroscedasticity robust standard errors in parentheses. ** denotes significance at the 5% level, * denotes significance at the 10% level.

Table 4C: Factor Endowments and Specialization in Services

GDP share	(1)	(2)	(3)	(4)
Obs	811	811	811	811
Years	1975-95	1975-95	1975-95	1975-95
Capital	0.001	-0.049**	-0.054**	-0.054**
	(0.0039)	*(0.0010)*	*(0.0113)*	*(0.0094)*
Population	0.029**	0.078**	0.083**	-0.076**
	(0.0057)	*(0.0106)*	*(0.0119)*	*(0.0296)*
Land	-0.038**	-0.048**	-0.048**	0.185**
	(0.0021)	*(0.0025)*	*(0.0027)*	*(0.0656)*
Sample	A	A	A	A
Specification	**(NULL)**	**(ALT1)**	**(ALT2)**	**(ALT3)**
Year dummies	yes	yes		
Country effects		yes		
Cty-year dummies			yes	yes
Region effects				yes
Prob>F(NULL-ALT)	N/A	0.0000	0.0000	0.0000
Prob>F(ALL)	0.0000	0.0000	0.0000	0.0000
R-squared	0.45	0.53	0.55	0.98
Sum of Coeff.	-0.0085	-0.0192	-0.019	0.055
Linear Homog	*(0.0091)*	*(0.0001)*	*(0.0002)*	*(0.4556)*
(p-value)	Reject	Reject	Reject	Accept
Maddala-Wu	*(0.1705)*	*(0.2460)*	*(0.0484)*	*(0.0412)*
(p-value)	Accept	Accept	Reject	Reject

Notes: Prob>F(NULL-ALT) is the p-value for an F-test of the null hypothesis that the coefficients on the variables excluded from specification (NULL) but incuded in the alternative specification are equal to 0. Prob>F(ALL) is the p-value for the conventional F-test that the coefficients on all independent variables are equal to zero. Sum of Coeff. is the sum of the estimated coefficients on factor endowments. Linear Homog. is the p-value for a test of the null hypothesis that the sum of the estimated coefficients on factor endowments is equal to zero. Maddala-Wu is the p-value for the Maddala and Wu (1999) panel data test of the null hypothesis that the residuals have a unit root. Huber-White heteroscedasticity robust standard errors in parentheses. ** denotes significance at the 5% level, * denotes significance at the 10% level.

Table 5: Factor Endowments and Specialization at the Aggregate Level

GDP share	(1)	(2)	(3)	(4)	(5)	(6)
Obs	696	696	696	696	696	696
Years	1975-95	1975-95	1975-95	1975-95	1975-95	1975-95
Capital	-0.016**	0.083**	-0.067**	0.012**	0.059**	-0.071**
	(0.0036)	*(0.0142)*	*(0.0129)*	*(0.0049)*	*(0.0116)*	*(0.0110)*
Low Educ	0.028**	-0.035*	0.007	-0.040**	0.015	0.024*
	(0.0049)	*(0.0181)*	*(0.0162)*	*(0.0066)*	*(0.0160)*	*(0.0148)*
Med Educ	-0.030**	0.077**	-0.048**	-0.014**	0.021**	-0.007
	(0.0046)	*(0.0188)*	*(0.0182)*	*(0.0028)*	*(0.0063)*	*(0.0053)*
High Educ	-0.001	-0.130**	0.131**	-0.011**	0.015**	-0.004
	(0.0033)	*(0.0160)*	*(0.0155)*	*(0.0031)*	*(0.0073)*	*(0.0058)*
Arable land	0.012**	0.020**	-0.032**	-0.001	0.017	-0.016
	(0.0009)	*(0.0025)*	*(0.0022)*	*(0.0021)*	*(0.0108)*	*(0.0111)*
Industry	Agric	Manu	Serv	Agric	Manu	Serv
Sample	B	B	B	B	B	B
Specification	**(ALT2)**	**(ALT2)**	**(ALT2)**	**(ALT3)**	**(ALT3)**	**(ALT3)**
Regional effects				yes	yes	yes
Cty-year dummies	yes	yes	yes	yes	yes	yes
Prob>F(NULL-ALT)	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Prob>F(ALL)	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
R-squared	0.67	0.50	0.62	0.96	0.97	0.98
Sum of Coeff.	-0.007	0.015	-0.009	-0.054	0.127	-0.074
Linear Homog	*(0.0000)*	*(0.0102)*	*(0.1273)*	*(0.0000)*	*(0.0000)*	*(0.0114)*
(p-value)	Reject	Reject	Accept	Reject	Reject	Reject
Maddala-Wu	*(0.0002)*	*(0.0106)*	*(0.0041)*	*(0.0000)*	*(0.2084)*	*(0.0092)*
(p-value)	Reject	Reject	Reject	Reject	Accept	Reject

Notes: Prob>F(NULL-ALT) is the p-value for an F-test of the null hypothesis that the coefficients on the variables excluded from specification (NULL) but incuded in the alternative specification are equal to 0. Prob>F(ALL) is the p-value for the conventional F-test that the coefficients on all independent variables are equal to zero. Sum of Coeff. is the sum of the estimated coefficients on factor endowments. Linear Homog. is the p-value for a test of the null hypothesis that the sum of the estimated coefficients on factor endowments is equal to zero. Maddala-Wu is the p-value for the Maddala and Wu (1999) panel data test of the null hypothesis that the residuals have a unit root. Huber-White heteroscedasticity robust standard errors in parentheses. ** denotes significance at the 5% level, * denotes significance at the 10% level.

Table 6: Factor Endowments and Specialization at the Aggregate Level (Long Differences)

△ GDP share	(1)	(2)	(3)
Obs	341	341	341
Years	1975-95	1975-95	1975-95
△ Capital	-0.006	0.062**	-0.057**
	(0.0043)	*(0.0104)*	*(0.0112)*
△ Low Educ	-0.035**	-0.009	0.044**
	(0.0043)	*(0.0185)*	*(0.0187)*
△ Med Educ	-0.031**	0.023**	0.008
	(0.0038)	*(0.0082)*	*(0.0081)*
△ High Educ	-0.010**	-0.002	0.012**
	(0.0017)	*(0.0042)*	*(0.0044)*
△ Arable Land	0.013**	0.018*	-0.031**
	(0.0035)	*(0.0108)*	*(0.0122)*
Industry	Agric	Manu	Serv
Sample	B	B	B
Specification	**(ALT4)**	**(ALT4)**	**(ALT4)**
Year dummies	yes	yes	yes
Difference period	10 years	10 years	10 years
Prob>F	0.0000	0.0000	0.0000
R-squared	0.46	0.21	0.19
Sum of Coeff.	-0.0685	0.0915	-0.0229
Linear Homog	*(0.0000)*	*(0.0016)*	*(0.4466)*
(p-value)	Reject	Reject	Accept

Notes: Sum of Coeff. is the sum of the estimated coefficients on factor endowments. Linear Homog. is the p-value for a test of the null hypothesis that the sum of the estimated coefficients on factor endowments is equal to zero. Huber-White heteroscedasticity robust standard errors in parentheses. ** denotes significance at the 5% level, * denotes significance at the 10% level.

Table 7: Average Shares of Sectors in GDP and Within-sample Average Absolute Prediction Errors

Country	Variable	(1) Agric	(2) Manu	(3) Serv
All countries *(a)*	GDP share	0.042	0.354	0.604
	Prediction Error 1 **(ALT2)**	0.582	0.133	0.068
	Prediction Error 2 **(ALT2 only cty-yr)**	6.316	0.499	0.159
	Prediction Error 3 **(ALT2 only endow.)**	6.985	0.501	1.099
Belgium *(a)*	GDP share	0.026	0.337	0.637
	Prediction Error 1 **(ALT2)**	0.519	0.056	0.038
	Prediction Error 2 **(ALT2 only cty-yr)**	5.151	0.572	0.107
	Prediction Error 3 **(ALT2 only endow.)**	5.632	0.476	1.099
Spain	GDP share	0.057	0.334	0.609
	Prediction Error 1 **(ALT2)**	1.529	0.170	0.072
	Prediction Error 2 **(ALT2 only cty-yr)**	10.603	0.439	0.159
	Prediction Error 3 **(ALT2 only endow.)**	10.277	0.524	1.095
France	GDP share	0.043	0.337	0.620
	Prediction Error 1 **(ALT2)**	0.677	0.118	0.046
	Prediction Error 2 **(ALT2 only cty-yr)**	8.700	0.526	0.124
	Prediction Error 3 **(ALT2 only endow.)**	9.267	0.461	1.104
Italy	GDP share	0.048	0.339	0.613
	Prediction Error 1 **(ALT2)**	0.280	0.121	0.053
	Prediction Error 2 **(ALT2 only cty-yr)**	2.940	0.502	0.151
	Prediction Error 3 **(ALT2 only endow.)**	3.822	0.471	1.133
Luxembourg	GDP share	0.024	0.342	0.634
	Prediction Error 1 **(ALT2)**	0 *(b)*	0 *(b)*	0 *(b)*
	Prediction Error 2 **(ALT2 only cty-yr)**	5.003	0.897	0.301
	Prediction Error 3 **(ALT2 only endow.)**	6.003	0.104	1.301
Netherlands	GDP share	0.047	0.353	0.600
	Prediction Error 1 **(ALT2)**	0.125	0.177	0.114
	Prediction Error 2 **(ALT2 only cty-yr)**	3.875	0.526	0.156
	Prediction Error 3 **(ALT2 only endow.)**	4.858	0.426	1.058
UK	GDP share	0.025	0.415	0.561
	Prediction Error 1 **(ALT2)**	0.462	0.145	0.098
	Prediction Error 2 **(ALT2 only cty-yr)**	6.634	0.444	0.213
	Prediction Error 3 **(ALT2 only endow.)**	7.755	0.650	1.057
Sample		B	B	B

Notes: table reports mean values for the whole sample and individual countries. Absolute proportional prediction errors are calculated as |s - s(P)| / s, where a capital P indicates a predicted value. Prediction error (ALT2) is based on the fitted values from specification (ALT2) using the disaggregated data on 5 factor endowments, and parameter estimates for this specification are reported in Table 5; Prediction error (ALT2, only cty-year) indicates that predicted values use the parameter estimates from specification (ALT2) but only the country-year dummies are used to construct predicted shares of GDP. Prediction error (ALT2, only endowments) indicates that predicted values use the parameter estimates from specification (ALT2) but only the 5 factor endowments are used to construct predicted shares of GDP.

(a) reported prediction errors exclude region Be1 (Brussels). Brussels is a capital city, and the share of Agriculture in this region is a clear outlier. As a robustness test, we re-estimated the model excluding this region; this produced very similar estimated coefficients to those reported earlier. (b) Luxembourg has only one NUTS-1 region. The fitted values for shares of sectors in GDP in the specification with country-year dummies are therefore exactly equal to the actual values (we estimate as many country-year coefficients as there are observations for Luxembourg). We experimented with treating Luxembourg as a region of Belgium; again this yielded very similar estimated coefficients to those reported earlier.

Table 8: Within-sample Average Absolute Prediction Errors over Time

Country	Variable	Period	(1) Agric	(2) Manu	(3) Serv
All countries *(a)*	Prediction Error **(ALT2)**	1980-85	0.566	0.131	0.070
	Prediction Error **(ALT2)**	1985-90	0.742	0.130	0.062
	Prediction Error **(ALT2)**	1990-95	0.560	0.131	0.054
Belgium *(a)*	Prediction Error **(ALT2)**	1980-85	0.403	0.031	0.029
	Prediction Error **(ALT2)**	1985-90	0.421	0.064	0.040
	Prediction Error **(ALT2)**	1990-95	0.708	0.078	0.048
Spain	Prediction Error **(ALT2)**	1980-85	1.195	0.191	0.083
	Prediction Error **(ALT2)**	1985-90	1.983	0.157	0.068
	Prediction Error **(ALT2)**	1990-95	1.513	0.164	0.064
France	Prediction Error **(ALT2)**	1980-85	0.724	0.111	0.042
	Prediction Error **(ALT2)**	1985-90	0.721	0.126	0.043
	Prediction Error **(ALT2)**	1990-94	0.565	0.117	0.043
Italy	Prediction Error **(ALT2)**	1980-85	0.340	0.107	0.047
	Prediction Error **(ALT2)**	1985-90	0.283	0.126	0.059
	Prediction Error **(ALT2)**	1990-95	0.213	0.129	0.053
Luxembourg	Prediction Error **(ALT2)**	1980-85	0 *(b)*	0 *(b)*	0 *(b)*
	Prediction Error **(ALT2)**	1985-90	0 *(b)*	0 *(b)*	0 *(b)*
Netherlands	Prediction Error **(ALT2)**	1980-85	0.174	0.224	0.170
	Prediction Error **(ALT2)**	1985-90	0.114	0.172	0.103
	Prediction Error **(ALT2)**	1990-95	0.083	0.141	0.065
UK	Prediction Error **(ALT2)**	1975-80	0.328	0.165	0.117
	Prediction Error **(ALT2)**	1980-85	0.501	0.129	0.082
Sample			B	B	B

Notes: table reports mean values for the whole sample and individual countries. Absolute proportional prediction errors are calculated as | *s - s(P)| / s* , where a capital P indicates a predicted value. Prediction error (ALT2) is based on the fitted values from specification (ALT2) using the disaggregated data on 5 factor endowments, and parameter estimates for this specification are reported in Table 5. *(a)* reported prediction errors exclude region Be1 (Brussels). Brussels is a capital city, and the share of Agriculture in this region is a clear outlier. As a robustness test, we re-estimated the model excluding this region; this produced very similar estimated coefficients to those reported earlier. *(b)* Luxembourg has only one NUTS-1 region. The fitted values for shares of sectors in GDP in the specification with country-year dummies are therefore exactly equal to the actual values (we estimate as many country-year coefficients as there are observations for Luxembourg). We experimented with treating Luxembourg as a region of Belgium; again this yielded very similar estimated coefficients to those reported earlier.

Table 9: Within-sample Average Absolute Prediction Errors in the Disaggregated Manufacturing Industries over Time

Country	Variable	Period	(1) Fuel	(2) Metal	(3) Mineral	(4) Chemical	(5) Machine	(6) Transport
All countries *(a),(b)*	Prediction Error **(ALT2)**	1980-85	0.575	2.109	0.269	0.444	0.453	0.587
	Prediction Error **(ALT2)**	1985-90	0.542	1.421	0.284	0.338	0.437	0.485
	Prediction Error **(ALT2)**	1990-95	0.524	1.251	0.289	0.389	0.501	0.519
Belgium *(a)*	Prediction Error **(ALT2)**	1980-85	0.290	0.373	0.600	0.168	0.266	1.043
	Prediction Error **(ALT2)**	1985-90	0.264	0.201	0.461	0.199	0.097	0.118
	Prediction Error **(ALT2)**	1990-95	0.366	0.149	0.365	0.178	0.155	0.138
Spain *(b)*	Prediction Error **(ALT2)**	1980-85	0.503	1.953	0.211	0.802	1.076	1.385
	Prediction Error **(ALT2)**	1985-90	0.379	2.289	0.282	0.477	0.905	0.637
	Prediction Error **(ALT2)**	1990-94	0.317	1.958	0.292	0.561	1.070	0.851
France	Prediction Error **(ALT2)**	1980-85	0.316	0.955	0.214	0.318	0.276	0.285
	Prediction Error **(ALT2)**	1985-90	0.348	1.078	0.256	0.339	0.250	0.368
	Prediction Error **(ALT2)**	1990-94	0.440	1.476	0.232	0.340	0.279	0.346
Italy	Prediction Error **(ALT2)**	1980-85	0.395	1.278	0.307	0.330	0.489	0.578
	Prediction Error **(ALT2)**	1985-90	0.390	1.162	0.282	0.308	0.449	0.675
	Prediction Error **(ALT2)**	1990-95	0.345	1.186	0.285	0.435	0.444	0.588
Luxembourg	Prediction Error **(ALT2)**	1980-85	0 *(c)*	0 *(c)*	0 *(c)*	0 *(c)*	0 *(c)*	0 *(c)*
	Prediction Error **(ALT2)**	1985-90	0 *(c)*	0 *(c)*	0 *(c)*	0 *(c)*	0 *(c)*	0 *(c)*
Netherlands	Prediction Error **(ALT2)**	1980-85	1.709	0.836	0.319	0.321	0.351	0.314
	Prediction Error **(ALT2)**	1985-90	1.637	0.707	0.376	0.248	0.407	0.352
	Prediction Error **(ALT2)**	1990-95	1.557	0.777	0.363	0.211	0.394	0.345
UK	Prediction Error **(ALT2)**	1975-80	0.631	2.906	0.460	0.553	0.329	0.495
	Prediction Error **(ALT2)**	1980-85	0.680	4.834	0.253	0.558	0.260	0.377
Sample			B	B	B	B	B	B

Country	Variable	Period	(1) Food	(2) Textile	(3) Paper	(4) Other	(5) Construction
All countries *(a)*	Prediction Error **(ALT2)**	1980-85	0.242	0.800	0.233	0.300	0.122
	Prediction Error **(ALT2)**	1985-90	0.241	0.714	0.311	0.293	0.111
	Prediction Error **(ALT2)**	1990-95	0.252	0.828	0.363	0.340	0.087
Belgium *(a)*	Prediction Error **(ALT2)**	1980-85	0.296	0.222	0.223	0.413	0.042
	Prediction Error **(ALT2)**	1985-90	0.233	0.171	0.185	0.167	0.052
	Prediction Error **(ALT2)**	1990-95	0.243	0.286	0.224	0.135	0.036
Spain	Prediction Error **(ALT2)**	1980-85	0.129	1.593	0.292	0.376	0.101
	Prediction Error **(ALT2)**	1985-90	0.145	0.828	0.335	0.436	0.108
	Prediction Error **(ALT2)**	1990-95	0.150	1.004	0.414	0.483	0.098
France	Prediction Error **(ALT2)**	1980-85	0.175	0.428	0.202	0.239	0.095
	Prediction Error **(ALT2)**	1985-90	0.162	0.458	0.201	0.284	0.076
	Prediction Error **(ALT2)**	1990-94	0.187	0.466	0.224	0.345	0.077
Italy	Prediction Error **(ALT2)**	1980-85	0.264	0.611	0.374	0.289	0.182
	Prediction Error **(ALT2)**	1985-90	0.290	0.782	0.544	0.273	0.152
	Prediction Error **(ALT2)**	1990-95	0.358	1.030	0.531	0.345	0.090
Luxembourg	Prediction Error **(ALT2)**	1980-85	0 *(c)*	0 *(c)*	0 *(c)*	0 *(c)*	0 *(c)*
	Prediction Error **(ALT2)**	1985-90	0 *(c)*	0 *(c)*	0 *(c)*	0 *(c)*	0 *(c)*
Netherlands	Prediction Error **(ALT2)**	1980-85	0.164	1.232	0.094	0.370	0.074
	Prediction Error **(ALT2)**	1985-90	0.267	1.226	0.133	0.247	0.102
	Prediction Error **(ALT2)**	1990-95	0.233	0.921	0.142	0.228	0.113
UK	Prediction Error **(ALT2)**	1975-80	0.404	1.142	0.295	0.333	0.095
	Prediction Error **(ALT2)**	1980-85	0.380	0.775	0.150	0.290	0.141
Sample			B	B	B	B	B

Notes: table reports mean values for the whole sample and individual countries. For full industry names, see Appendix A. Absolute proportional prediction errors are calculated as |s - s(P)| / s, where a capital P indicates a predicted value. Prediction error (ALT2) is based on the fitted values from specification (ALT2) using the disaggregated data on 5 factor endowments, and parameter estimates for this specification are reported in Table 5. *(a)* the reported prediction errors exclude region Be1 (Brussels). Brussels is a capital city, and the shares of some disaggregated manufacturing industries in this region are clear outliers. As a robustness test, we re-estimated the model excluding this region; this produced very similar estimated coefficients to those reported earlier.

(b) the reported prediction errors exclude the Metal industry in region Esp7 (Canaries). Metal constitutes a very small share of GDP in this region. As a robustness test, we re-estimated the model excluding this region; this produced very similar estimated coefficients to those reported earlier. (c) Luxembourg has only one NUTS-1 region. The fitted values for shares of sectors in GDP in the specification with country-year dummies are therefore exactly equal to the actual values (we estimate as many country-year coefficients as there are observations for Luxembourg). We experimented with treating Luxembourg as a region of Belgium; again this yielded very similar estimated coefficients to those reported earlier.

7 Appendix A

Table 1: Sample Composition

Country	Sample A	Sample B	Number of NUTS1 regions
Belgium	1975-95	1979-95	3 (be1-be3)
Spain	1980-95	1980-94	7 (esp1- esp7)
France	1975-95	1977-94	8 (fra1-fra8)
Italy	1975-95	1980-95	11 (ita1-ita9, itaa/b)
Luxembourg	1975-95	1979-90	1 (lux)
Netherlands	1975-95	1977-95	4 (ndl1-ndl4)
United Kingdom	1975-86	1975-86	11 (uk1-uk9, uka/b)

Table 2: Industry Composition

Code	Industry Description
Aggregate Industries	
1	Agricultural Sector: Food, Forestry and Fishery Products **(Agric)**
2	Manufacturing Sector **(Manu)**
3	Services Sector: Market Services **(Serv)**
Disaggregated Manufacturing Industries	
4	Fuel And Power Products **(Fuel)**
5	Ferrous And Non-Ferrous Ores And Metals, Other Than Radioactive **(Metal)**
6	Non-Metallic Minerals And Mineral Products **(Mineral)**
7	Chemical Products **(Chem)**
8	Metal Products, Machinery, Equipment And Electrical Goods **(Machine)**
9	Transport Equipment **(Transp)**
11	Food, Beverages And Tobacco **(Food)**
12	Textiles And Clothing, Leather And Footwear **(Textile)**
13	Paper And Printing Products **(Paper)**
14	Products Of Various Industries **(Other)**
15	Building And Construction **(Constr)**

Table 3: Regions Included in the Sample

Code	Description
Belgium	
Be1	Brussels
Be2	Vlaams Gewest
Be3	Region Wallonne
Spain	
Esp1	North West (E)
Esp2	North East (E)
Esp3	Madrid
Esp4	Centre (E)
Esp5	East (E)
Esp6	South (E)
Esp7	Canaries
France	
Fra1	Ile De France
Fra2	Bassin Parisien
Fra3	Nord-Pas-de-Calais
Fra4	East (F)
Fra5	West (F)
Fra6	South West (F)
Fra7	Centre-East (F)
Fra8	Mediterranean
Italy	
Ita1	North West (I)
Ita2	Lombardia
Ita3	Nord East (I)
Ita4	Emilia-Romagna
Ita5	Centre (I)
Ita6	Lazio
Ita7	Abruzzo-Molise
Ita8	Campania
Ita9	South (I)
Itaa	Sicily
Itab	Sardinia
Luxembourg	
Lux	Luxembourg (Grand-Duche)
Netherlands	
Nld1	North-Netherland
Nld2	East-Netherland
Nld3	West-Netherland
Nld4	South-Netherland

Code	Description
United Kingdom	
UK1	North (UK)
UK2	Yorkshire And Humberside
UK3	East Midlands
UK4	East Anglia
UK5	South East (UK)
UK6	South West (UK)
UK7	West Midlands
UK8	North West (UK)
UK9	Wales
UKA	Scotland
UKB	Northern Ireland

Appendix B

B1. Regional-level Data on Production and Endowments

1. Value Added: current price value-added, millions of ECUs, from Regio dataset, Eurostat.
2. GDP: current price, millions of ECUs, from Regio dataset, Eurostat.
3. Population: total population, thousands of people, from Regio dataset, Eurostat.
4. Land: total land area, thousands of hectares, from Regio dataset, Eurostat.
5. Arable Land: total arable land area, thousands of hectares, from Regio dataset, Eurostat.
6. Capital Stock: constructed by the perpetual inventory method (see, for example, Barro and Sala-i-Martin 1995) using regional-level investment data (Gross Fixed Capital Formation), constant 1990 prices, millions of ECUs. The main source for the investment data is the Regio dataset, Eurostat. Current price investment was converted into constant prices using price deflators from the Penn World Tables, 5.6. For some countries, regional current price investment data were extended backwards in time using country-level information from the IMF International Financial Statistics.

B2. Summary of Educational Attainment Data Sources

Following the labour market literature (see, for example, Nickell and Bell 1996 and Machin and Van Reenen 1998), educational attainment is grouped into three categories: low, medium and high. 'Low education' is no or primary education, while 'high education' is College degree or equivalent. 'Medium education' corresponds to all intermediate levels of educational attainment, including secondary school and vocational qualifications. Using individual country labour force surveys, we compute the percentage of the population with each level of educational attainment. The endowment variables included in the regressions are these percentages multiplied by the population data from Regio, Eurostat.

1. Belgium: regional data on educational attainment from *Annuaire de Statistiques Regionales*. Years available are 1970, 1977, 1981 and 1991. Linear interpolation of the data.

2. Spain: educational attainment data from Spanish Labour Force, *Instituto Nacional de Estadistica*. Years available are 1977, 1979, 1981, and 1983-94. Linear interpolation of the data when required.

3. Italy: educational attainment data from 1986-97 is from *Forze di Lavoro* and *Rilevazione delle forze di Lavore*, ISTAT. For years prior to 1986, the regional data is extended backwards in time using country-level information from Nickell *et al.* (2000).

4. France: educational attainment data from *Key data on Education*, DG for Education and Culture, European Commission. Years available are 1993 and 1995. Linear interpolation of the data for 1994. The regional data are extended backwards in time country-level information from Nickell *et al.* (2000).

5. Netherlands: Data from *National Statistical Office*, years 1992-98. The regional data are extended backwards in time using country-level information from Nickell *et al.* (2000).

6. Luxembourg: Data are from Belgian region closest to Luxembourg (be3, Region Wallone).

7. United Kingdom: Data from the Labour Force Survey, years 1977, 1979, 1981, and 1983-94. Linear interpolation of the data when required. Bibliographic citation: Office for National Statistics Labour Market Statistics Group, Department of Finance and Personnel (Northern Ireland), Central Survey Unit, Quarterly Labour Force Survey. Data distributed by the Data Archive, Colchester, Essex. Data disclaimer: although all efforts are made to ensure the quality of the materials, neither the copyright holder, the original data producer, the relevant funding agency, The Data Archive, bear any responsibility for the accuracy or comprehensiveness of these materials.

Appendix C

Table C1A: Factor Endowments and Specialisation at the Disaggregate Level

GDP share	(1)	(2)	(3)	(4)	(5)	(6)
Obs	696	689	689	696	696	693
Years	1975-95	1975-95	1975-95	1975-95	1975-95	1975-95
Capital	-0.015*	-0.014**	-0.003**	0.007**	0.057**	0.005*
	(0.0081)	*(0.0025)*	*(0.0012)*	*(0.0017)*	*(0.0058)*	*(0.0026)*
Low Educ	-0.001	0.011**	0.002*	0.003	-0.051**	-0.008**
	(0.0108)	*(0.0033)*	*(0.0013)*	*(0.0021)*	*(0.0088)*	*(0.0030)*
Med Educ	-0.013	0.018**	0.005**	-0.001	0.049**	0.022**
	(0.0116)	*(0.0041)*	*(0.0016)*	*(0.0032)*	*(0.0125)*	*(0.0045)*
High Educ	0.002	-0.013**	-0.008**	-0.002	-0.041**	-0.012**
	(0.0065)	*(0.0025)*	*(0.0012)*	*(0.0022)*	*(0.0071)*	*(0.0033)*
Arable Land	0.007**	0.001	0.002**	-0.001	0.002**	-0.001
	(0.0016)	*(0.0005)*	*(0.0002)*	*(0.0004)*	*(0.0010)*	*(0.0004)*
Industry	Fuel	Metal	Mineral	Chem	Machine	Transp
Sample	B	B	B	B	B	B
Specification	**(ALT2)**	**(ALT2)**	**(ALT2)**	**(ALT2)**	**(ALT2)**	**(ALT2)**
Cty-year dummies	yes	yes	yes	yes	yes	yes
Prob>F(NULL-ALT2)	0.0162	0.0000	0.0000	0.0000	0.0000	0.0000
Prob>F(ALL)	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
R-squared	0.27	0.68	0.50	0.33	0.39	0.36
Sum of Coeff.	-0.0191	0.0030	-0.0020	0.0060	0.0160	0.0059
Linear Homog	*(0.0000)*	*(0.0093)*	*(0.0060)*	*(0.0000)*	*(0.0000)*	*(0.0000)*
(p-value)	Reject	Reject	Reject	Reject	Reject	Reject
Maddala-Wu	*(0.0101)*	*(0.0000)*	*(0.0000)*	*(0.1195)*	*(0.0538)*	*(0.0000)*
(p-value)	Reject	Reject	Reject	Accept	Reject	Reject

Notes: for full industry names, see Appendix A. Prob>F(NULL-ALT) is the p-value for an F-test of the null hypothesis that the coefficients on the variables excluded from specification (NULL) but incuded in the alternative specification are equal to 0. Prob>F(ALL) is the p-value for the conventional F-test that the coefficients on all independent variables are equal to zero. Sum of Coeff. is the sum of the estimated coefficients on factor endowments. Linear Homog. is the p-value for a test of the null hypothesis that the sum of the estimated coefficients on factor endowments is equal to zero. Maddala-Wu is the p-value for the Maddala and Wu (1999) panel data test of the null hypothesis that the residuals have a unit root. Huber-White heteroscedasticity robust standard errors in parentheses. ** denotes significance at the 5% level, * denotes significance at the 10% level.

Table C1B: Factor Endowments and Specialisation at the Disaggregate Level

GDP share	(1)	(2)	(3)	(4)	(5)
Obs	696	696	696	696	696
Years	1975-95	1975-95	1975-95	1975-95	1975-95
Capital	0.012**	0.022**	0.015**	0.021**	-0.022**
	(0.0030)	*(0.0031)*	*(0.0009)*	*(0.0017)*	(0.0024)
Low Educ	0.012**	0.003	-0.018**	-0.014**	0.027**
	(0.0050)	*(0.0040)*	*(0.0015)*	*(0.0024)*	(0.0035)
Med Educ	-0.022**	0.018**	0.011**	0.017**	-0.031**
	(0.0071)	*(0.0046)*	*(0.0016)*	*(0.0027)*	*(0.0043)*
High Educ	-0.011**	-0.035**	-0.005**	-0.021**	0.016**
	(0.0034)	*(0.0039)*	*(0.0010)*	*(0.0019)*	(0.0027)
Arable Land	0.004**	0.001	-0.001	0.003**	0.003**
	(0.0005)	*(0.0005)*	*(0.0002)*	*(0.0003)*	*(0.0004)*
Industry	Food	Textile	Paper	Other	Constr
Sample	B	B	B	B	B
Specification	**(ALT2)**	**(ALT2)**	**(ALT2)**	**(ALT2)**	**(ALT2)**
Cty-year dummies	yes	yes	yes	yes	yes
Prob>F(NULL-ALT2)	0.0000	0.0000	0.0000	0.0000	0.0000
Prob>F(ALL)	0.0000	0.0000	0.0000	0.0000	0.0000
R-squared	0.57	0.43	0.64	0.54	0.69
Sum of Coeff.	-0.0051	0.0085	0.0027	0.0059	-0.0074
Linear Homog	*(0.0007)*	*(0.0000)*	*(0.0000)*	*(0.0000)*	*(0.0000)*
(p-value)	Reject	Reject	Reject	Reject	Reject
Maddala-Wu	*(0.2104)*	*(0.0068)*	*(0.0017)*	*(0.0028)*	*(0.1405)*
(p-value)	Accept	Reject	Reject	Reject	Accept

Notes: for full industry names, see Appendix A. Prob>F(NULL-ALT) is the p-value for an F-test of the null hypothesis that the coefficients on the variables excluded from specification (NULL) but incuded in the alternative specification are equal to 0. Prob>F(ALL) is the p-value for the conventional F-test that the coefficients on all independent variables are equal to zero. Sum of Coeff. is the sum of the estimated coefficients on factor endowments. Linear Homog. is the p-value for a test of the null hypothesis that the sum of the estimated coefficients on factor endowments is equal to zero. Maddala-Wu is the p-value for the Maddala and Wu (1999) panel data test of the null hypothesis that the residuals have a unit root. Huber-White heteroscedasticity robust standard errors in parentheses. ** denotes significance at the 5% level, * denotes significance at the 10% level.