Factor Endowments and Production in European Regions

Stephen Redding and **Mercedes Vera-Martin**

London School of Economics and CEPR, London;
International Monetary Fund, Washington, D.C.

Abstract: This paper analyses patterns of production across 14 industries in 45 regions from 7 European countries since 1975. We estimate an equation from neoclassical trade theory that relates an industry's share of a region's GDP to factor endowments, relative prices and technology. The strict version of the Heckscher–Ohlin model that assumes identical relative prices and technology is rejected against more general alternatives. However, factor endowments play a statistically significant and quantitatively important role in explaining production patterns. Factor endowments are more successful at explaining patterns of production in aggregate industries (Agriculture, Manufacturing and Services) than in disaggregated industries within manufacturing. JEL no. F11, F14, R13
Keywords: Factor endowments; Heckscher–Ohlin; neoclassical model; Europe

1 Introduction

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 the empirical validity of its assumptions of identical and homothetic prefer-

Remark: This paper is produced as part of the Globalisation Programme of the ESRC funded Centre for Economic Performance and is a revised version of CEPR Discussion Paper 3755. The views expressed are those of the authors and not necessarily those of any institution with which the authors are affiliated. Vera-Martin's research was funded by a Central Bank of Spain Scholarship. We are grateful to an anonymous referee and seminar participants at the London School of Economics and Yale University for helpful comments. We would like to thank Maia Guell-Rotllan, Steve Machin, Marco Manacorda, Steve Nickell, Jan van Ours, and Jo Swaffield for help with the data. We are also grateful to Giorgia Albertin and Estela Montado for research assistance. Please address correspondence to Stephen Redding, Department of Economics, London School of Economics, Houghton Street, London. WC2A 2AE United Kingdom; e-mail: s.j.redding@lse.ac.uk

 DOI: 10.1007/s10290-006-0055-y

ences, identical technologies, and no barriers to trade. This paper examines the ability of the HO model to explain production patterns at the regional level in Europe using a newly constructed panel data set on output in 14 industries and endowments of 5 factors of production for 45 NUTS-1 regions from 7 European countries since 1975.[1] 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 European Union provides an interesting context within which to explore the relationship between factor endowments and production. The ongoing process of economic integration introduces exogenous variation in relative prices. We control for this variation and examine whether the relationship between endowments and production within countries has been strengthened or weakened by deeper integration.

Much existing empirical work on the international location of production has, for reasons of data availability, been concerned with the manufacturing sector. We consider both manufacturing and the non-manufacturing sectors that account for more than 70 per cent of GDP in many NUTS-1 regions. We analyse the determinants of production structure for three broad industries (Agriculture, Manufacturing and Services) as well as for 11 more finely detailed industries within manufacturing (including, for example, Textiles and Chemicals). We consider five factor endowments: high-education, medium-education, and low-education individuals, physical capital, and land area. The analysis focuses on production rather than trade because the central predictions of the HO model are for producer equilibrium and, by focusing on production, we avoid making restrictive assumptions about consumer behaviour.[2]

While the use of regional data has many advantages, it means that the standard trade assumption that factor endowments are exogenous and perfectly immobile across locations is unlikely to apply. We show that the general equilibrium relationships between production structure and factor endowments that we derive under the null hypothesis of the HO model

---

[1] 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.

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

and under more general alternatives hold irrespective of the degree of factor mobility.[3]

Factor mobility does, however, change the interpretation of these relationships. If factor endowments are geographically immobile, the general equilibrium relationship between production structure and factor endowments should be interpreted in *supply-side* terms. Exogenous changes in factor endowments cause endogenous changes in production structure (production moves in response to factor endowments). If factor endowments are geographically mobile, there is also a *demand-side* interpretation, whereby exogenous changes in demand and hence production cause factor endowments to move endogenously across regions (factor endowments move in response to production structure). The equations that we estimate incorporate both demand- and supply-side effects. We estimate the general equilibrium relationship between production structure and factor endowments that these forces imply, and we test whether or not this relationship takes the restrictive form implied by the HO model.

Our main empirical findings are as follows. First, the strict version of the Heckscher–Ohlin model that assumes identical relative prices and technology is rejected against more general alternative hypotheses that allow for variation in relative prices, technology and other omitted variables such as market access. Second, factor endowments are nonetheless statistically significant and quantitatively important determinants of variation in production structure across European regions.

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.

[3] 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 (1998) for evidence on the United Kingdom.

Fourth, factor endowments are more successful in explaining patterns of production at the aggregate level in Agriculture, Manufacturing and Services than in disaggregated manufacturing industries. Since there are a greater number of goods relative to factors at the disaggregated level, this finding is consistent with production indeterminacy in the HO model with more goods than factors. It is also consistent with the idea that variation in relative prices, technology and other variables omitted from the HO model may be particularly great in individual disaggregated manufacturing industries.

The paper is related to two main strands of existing literature. First, there are cross-country studies of the relationship between factor endowments, production and the net factor content of trade, including among others Bowen et al. (1987), Gabaix (2005), Harrigan (1995, 1997), Kohli (1991), Leamer (1984), Nickell et al. (2001), Redding (2002), Schott (2003) and Trefler (1995). These literatures find evidence of cross-country technology differences and non-factor price equalization. Only when the HO model is augmented to allow for these and other considerations, does it provide a satisfactory explanation of the data.

Second, a related literature examines the predictions of the HO model at the regional level. Hanson and Slaughter (2002) and Gandal et al. (2004) examine how immigration affects relative factor endowments and hence output mix across US states and in Israel. Davis et al. (1997) present evidence that the HO model's assumption of factor price equalization and its predictions for net trade in factor services are rejected using cross-country data but confirmed using Japanese regional data. Bernstein and Weinstein (2002) find that regressions of output on factor endowments yield prediction errors that are substantially larger using Japanese regional data than using cross-country data. This is consistent with the idea that there is production indeterminacy, which is likely to be larger for the lower values of trade costs observed across regions within a country.

The main contributions of this paper are the use of regional data on Europe, where there has been little work testing neoclassical trade theory, combined with an estimation specification derived from the translog revenue function of the neoclassical model, which enables us to test the HO null against more general alternatives. Unlike many of the above studies, our data allow us to consider both manufacturing and non-manufacturing industries, and to examine the performance of the HO null for aggregate and disaggregate industries. Our findings shed light on the statistical significance and quantitative importance of factor endowments in explaining production structure across regions within Europe for both broad and finely detailed industries.

The plan of attack is as follows. Section 2 introduces the theoretical framework and derives predictions for production patterns under the HO null and under more general alternatives. 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

2.1 Neoclassical Theory of Trade and Production

The theoretical framework is the neoclassical theory of trade and production (see, in particular, 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 occurs under perfect competition and constant returns to scale.[4]

The neoclassical model allows for regional differences in factor endowments as well as region-industry differences in technology and relative prices. The HO model is a special case of the neoclassical model where all regions have identical relative prices and technology.

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}.[5]

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

[5] 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.

We allow for Hicks-neutral region-industry-time technology differences so that the production function 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 and v_{zjt} denotes the vector of factors employed in industry j of region z at time t. In this case, the revenue 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}.[6] Changes in technology in industry j of region z have analogous effects on revenue to changes in industry j prices.

We assume a translog revenue function. This flexible functional form provides an arbitrarily close local approximation to the true underlying revenue function:

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

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:

$$\begin{aligned} s_{zjt} &\equiv \frac{p_{zjt} y_{zjt}(p_{zt}, v_{zt})}{r(p_{zt}, v_{zt})} \\ &= \beta_{0j} + \sum_k \beta_{jk} \ln p_{zkt} + \sum_k \beta_{jk} \ln \theta_{zkt} + \sum_i \gamma_{ji} \ln v_{zit} . \end{aligned} \tag{2}$$

Equation (2) is a general equilibrium relationship between the share of an industry in regional production on the left-hand side and relative prices, technology and factor endowments on the right-hand side. A change in the

[6] Cross-country 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. See Dixit and Norman (1980) for further discussion.

right-hand side variables (e.g. in relative prices) will lead to adjustments in industrial structure until a new equilibrium consistent with producer optimization, consumer optimization, goods market clearing and factor market clearing is attained.

The direction and size of the adjustment in the output of individual industries is revealed by the coefficients on relative prices, technology and factor endowments which vary across industries j. In particular, the coefficients on factor endowments i for each industry j, γ_{ji}, contain information on the factor intensity of industries since they relate to the cross-derivatives of the revenue function with respect to price and factor endowments.[7]

The translog revenue function implies coefficients on relative prices, technology and factor endowments that are constant across regions and over time. This is true even without factor price equalization. Indeed, where there are regional differences in prices and technology, factor price equalization will not be observed in general. 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 (2).

The analysis makes no assumptions about whether regions are large or small, and the analysis 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 perfectly immobile, perfectly mobile or exhibit limited mobility across regions. In each case, (2) holds.

2.2 Heckscher–Ohlin Null Hypothesis

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) are perfectly colinear with

---

[7] With many goods, many factors of production, and at least as many factors as goods ($M \geq N$), the natural concept of factor intensity is the cross-derivative $\partial^2 r(p, v)/\partial p_j \partial v_i$ (see Dixit and Norman 1980: 54–59). Under a translog revenue function, the γ_{ji} are proportional to, and take the same sign as, these measures of factor intensity ($\gamma_{ji}/(p_j v_i) = \partial^2 r(p, v)/\partial p_j \partial v_i$). The relationship between cross-derivatives of the revenue function and factor intensity is particularly clear when there are equal numbers of goods and factors ($M = N$). In this case, the cross-derivatives of the revenue function equal the relevant element of the inverse of the matrix of unit factor input requirements ($y = A^{-1} v$ and $y = Rv$, where R is the matrix of cross-derivatives).

a full set of time dummies. Therefore, the relative price and technology terms may be replaced by time dummies, which have industry-specific coefficients to reflect the differential impact of changes in relative prices and technology on production across industries:

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

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 (2) has been absorbed into the ϕ_{jt}.

Under the HO model's identifying assumptions of identical relative prices and technology, the general equilibrium relationship between factor endowments and production structure (γ_{ji} for all factors i and industries j) may be consistently estimated using (3).

2.3 Alternative Hypotheses

Unlike the Heckscher–Ohlin model, neoclassical theory allows differences in relative prices and technology to be the basis for comparative advantage in addition to variation in factor endowments. Data on relative prices and technology within individual industries are not available across a broad cross-section of European regions. Instead, we consider a series of progressively more general econometric models of relative prices and technology, which when substituted into (2) provide alternatives to the HO null. In practice there may also be other determinants of patterns of production besides relative prices and technology, such as institutions and market access, which our more general models also control for.

First, we model relative price and technology differences with a country-industry fixed effect (η_{cj}), industry-time dummies ($\phi_{jt} d_t$) and a stochastic error, which yields our first alternative hypothesis:

$$\text{(ALT1)} \qquad s_{zjt} = \sum_i \gamma_{ji} \ln v_{zit} + \sum_t \phi_{jt} d_t + \eta_{cj} + \varepsilon_{zjt} \,, \qquad (4)$$

where the constant β_{0j} has again been absorbed into the other coefficients.

This specification differs from the null hypothesis through the inclusion of the country-industry fixed effects, which allow for permanent cross-country differences in relative prices and technology that have non-neutral effects on industrial structure in general equilibrium, as well as other unobserved time-invariant determinants of production structure that vary across countries. A test of the statistical significance of the country-industry fixed effects provides a test of the HO null hypothesis against the more general alternative.

Second, we model relative price and technology differences with country-industry-time dummies ($\mu_{cjt} d_{ct}$) and a stochastic error, which yields our second alternative hypothesis:

$$\text{(ALT2)} \qquad s_{zjt} = \sum_i \gamma_{ji} \ln v_{zit} + \sum_t \phi_{jt} d_t + \sum_{c \neq c'} \sum_t \mu_{cjt} d_{ct} + \varepsilon_{zjt}\,, \tag{5}$$

where d_{ct} are {0, 1} country-time dummies, μ_{cjt} are the industry-specific coefficients on these country-time dummies, and the constant β_{0j} has again been absorbed into other coefficients. For ease of comparability with the null hypothesis, we continue to include the year dummies ($\phi_{jt} d_t$), and exclude country c' from the country-year dummies (d_{ct}) so that the estimated coefficients (μ_{cjt}) are conditional means relative to country c'.

The country-year dummies (d_{ct}) allow for differential changes in relative prices and technology across countries and over time. The industry-specific coefficients (μ_{cjt}) capture the uneven effects of these changes in relative prices and technology across industries in general equilibrium. The country-year dummies control for a variety of determinants of production patterns. They capture country-specific changes in relative prices associated with the ongoing process of European integration. They also control for violations in the law of one price across countries due, for example, to border effects as well as other unobserved country-time varying determinants of production structure. A test of the statistical significance of the country-year dummies provides a test of the HO null hypothesis against the more general alternative.

Third, we extend the model of relative prices and technology further to allow for a region-industry fixed effect (η_{zj}) in addition to the country-time dummies and a stochastic error, which yields our third alternative hypothesis:

$$\text{(ALT3)} \qquad s_{zjt} = \sum_i \gamma_{ji} \ln v_{zit} + \sum_t \phi_{jt} d_t + \sum_{c \neq c'} \sum_t \mu_{cjt} d_{ct} + \eta_{zj} + \varepsilon_{zjt}\,, \tag{6}$$

where the constant β_{0j} has again been absorbed into the other coefficients and we adopt the same normalization of the country-year dummies.

The region-industry fixed effects allow for permanent differences in relative prices and technology across regions and industries that influence production structure in general equilibrium, as well as other unobserved region-specific time-invariant determinants of production structure. A test of the joint statistical significance of the country-year dummies and the region-industry fixed effects provides a test of the HO null hypothesis against the more general alternative.

3 Data Description

The main source of data is the Regio data set compiled by the European Statistics Office (Eurostat). We analyse 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, the Netherlands, Spain and the United Kingdom.[8] 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 (e.g. Belgium and France) and others located further towards the 'periphery' (e.g. 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 United Kingdom, 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 in Appendix A. We show below that there is also substantial variation in specialization across NUTS-1 regions within a country, for example from the North of Italy to Sicily.

Patterns of production are analysed 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.

---

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

The Regio data set provides information on current-price 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.[9] 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.[10]

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 1975–1995 (Sample A). Second, for the disaggregated manufacturing industries and for the five 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 per cent in Be1 (Brussels) to 11.86 per cent in Esp4 (Centre), while the GDP share of Services in the same year ranges from 81.61 per cent in Be1 (Brussels) to 49.57 per cent 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 per cent in 1980 to 5.39 per cent in 1995, while the GDP share of

---

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

10 See Appendix A for further information concerning the data used.

Table 1: *Shares of Agriculture, Manufacturing and Services in GDP (per cent)*

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
		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
		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

Table 1: *continued*

Region	Year	Agric	Manu	Serv	Region	Year	Agric	Manu	Serv
Ita9	1975	14.17	31.02	54.81	Uk4	1975	6.56	36.12	57.32
	1985	9.42	27.81	62.77		1980	5.78	35.93	58.29
	1995	6.49	24.49	69.02		1985	3.07	36.02	60.91
Itaa	1975	12.56	29.54	57.90	Uk5	1975	0.82	32.14	67.03
	1985	9.36	28.33	62.32		1980	0.85	32.51	66.65
	1995	5.73	21.47	72.80		1985	0.48	29.46	70.06
Itab	1975	9.09	36.40	54.51	Uk6	1975	3.27	36.86	59.87
	1985	5.96	33.64	60.39		1980	3.23	35.60	61.17
	1995	4.11	27.37	68.52		1985	2.16	34.36	63.48
					Uk7	1975	1.51	48.94	49.55
		Luxembourg				1980	1.71	46.34	51.94
Lux	1975	3.24	39.18	57.57		1985	1.23	43.64	55.13
	1985	2.36	34.04	63.60	Uk8	1975	0.87	45.53	53.60
	1995	1.21	31.61	67.19		1980	0.70	45.46	53.84
						1985	0.53	44.32	55.15
		United Kingdom			Uk9	1975	2.81	45.57	51.61
Uk1	1975	2.04	50.04	47.93		1980	2.88	45.14	51.98
	1980	1.58	47.29	51.12		1985	2.48	45.93	51.59
	1985	1.37	41.37	57.25	Uka	1975	2.98	43.73	53.28
Uk2	1975	2.31	46.46	51.23		1980	2.39	42.17	55.44
	1980	1.92	44.88	53.21		1985	1.69	38.82	59.49
	1985	1.43	41.40	57.17	Ukb	1975	3.40	43.37	53.23
Uk3	1975	2.76	48.36	48.88		1980	3.30	37.21	59.49
	1980	2.76	46.45	50.79		1985	2.89	36.00	61.11
	1985	1.59	44.29	54.11					

Note: Figures may not sum exactly to 100 due to rounding.

Services in Fra3 (Nord-Pas-de-Calais) rises from 46.68 per cent in 1975 to 67.00 per cent 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. 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 per cent between 1980 and 1994, while the share of Paper in Fra3 (Nord-Pas-de-Calais) rises by 24 per cent over the same period.

4 Econometric Specification

Our null and alternative specifications are derived from the theory, as explained above, and are reproduced below:

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

$$\text{(ALT1)} \qquad s_{zjt} = \sum_i \gamma_{ji} \ln v_{zit} + \sum_t \phi_{jt} d_t + \eta_{cj} + \varepsilon_{zjt}$$

$$\text{(ALT2)} \qquad s_{zjt} = \sum_i \gamma_{ji} \ln v_{zit} + \sum_t \phi_{jt} d_t + \sum_{c \neq c'} \sum_t \mu_{cjt} d_{ct} + \varepsilon_{zjt}$$

$$\text{(ALT3)} \qquad s_{zjt} = \sum_i \gamma_{ji} \ln v_{zit} + \sum_t \phi_{jt} d_t + \sum_{c \neq c'} \sum_t \mu_{cjt} d_{ct} + \eta_{zj} + \varepsilon_{zjt} \, .$$

Since all coefficients vary across industries, the specifications are estimated separately for each industry, pooling observations across regions and over time. Standard errors reported below are heteroscedasticity-robust and adjusted for clustering on country-year.

As we start from (NULL) and move to (ALT3), we consider progressively more general models. However, in moving 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 analyse the results of 10-year difference estimators. We thus obtain a fourth alternative specification:

$$\text{(ALT4)} \qquad \triangle_{10} s_{zjt} = \sum_i \gamma_{ji} \triangle_{10} \ln v_{zit} + \sum_t \zeta_{jt} d_t + \psi_{zjt} \, , \qquad (7)$$

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 including time dummies that are common across countries together with their industry-specific coefficients.

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 (1975–1995) 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-level 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 recent advances in panel data unit root tests, exploiting the Maddala and Wu (1999) methodology to test for the stationarity of the residuals.[11]

Second, neoclassical trade theory assumes that the production technology is constant returns to scale, implying that the revenue function in equation (1) is homogenous of degree one in factor endowments, and implying that the industry share equation (2) is homogenous of degree zero in factor endowments. Therefore, a test of the null hypothesis that the sum of the estimated coefficients on factor endowments is equal to zero in specifications (NULL) and (ALT1)–(ALT4), $\sum_i \gamma_{ji} = 0$, provides another model specification test.

[11] 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. (2003), Levin and Lin (1992), Pedroni (1999), Pesaran et al. (1999) and Quah (1994).

5 Empirical Results

5.1 Aggregate Industries, Aggregate Endowments

We begin in column (1) of Table 2 by reporting the results of estimating the HO specification (NULL) for the aggregate industries (Agriculture, Manufacturing and Services) using our three broad measures of factor endowments (population, physical capital and land area). We find a relationship between regional patterns of production and factor endowments which is statistically significant at the 5 per cent level and, from the regression R^2, the HO specification explains some 30–45 per cent of the variation in production patterns across European regions.

In column (2) of Table 2 (panels A–C), we relax the assumption of identical relative prices and technology, and move to specification (ALT1). The country-industry fixed effects, which are excluded from the null hypothesis, are highly statistically significant as shown in the first of the F-statistics reported towards the bottom of the column. 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.

Therefore, the strict version of the Heckscher–Ohlin model that assumes identical relative prices and technology is rejected against more general alternatives. The country-industry fixed effects are not only statistically significant but also quantitatively important, as reflected in the rise in the regression R^2. The change in the pattern of estimated coefficients when moving from (NULL) to (ALT1) indicates that these additional controls are important for identifying the nature of the relationship between production structure and factor endowments.

In column (3) of Table 2 (panels A–C), we consider the more general alternative specification (ALT2). We allow for differential movements across countries and over time in relative prices, technology and other variables omitted from the HO model by including a set of country-year dummies in each industry regression for all countries except the base country c'. From the first of the F-statistics reported towards the bottom of this column, the coefficients on the country-year dummies are highly statistically significant

Table 2: *Factor Endowments and Specialization in Three Aggregate Sectors*

	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
	811 obs.	811 obs.	811 obs.	811 obs.	811 obs.	811 obs.	811 obs.	811 obs.	811 obs.	811 obs.	811 obs.	811 obs.
	1975-95	1975-95	1975-95	1975-95	1975-95	1975-95	1975-95	1975-95	1975-95	1975-95	1975-95	1975-95
	A. *Agriculture*				B. *Manufacturing*				C. *Services*			
Capital	0.004*	−0.022**	−0.019**	0.011**	−0.005	0.071**	0.073**	0.043**	0.001	−0.049**	−0.054**	−0.054**
	(0.0024)	*(0.0050)*	*(0.0055)*	*(0.0045)*	*(0.0058)*	*(0.0205)*	*(0.0241)*	*(0.0080)*	*(0.0045)*	*(0.0163)*	*(0.0196)*	*(0.0097)*
Population	−0.021**	0.002	−0.002	−0.129**	−0.008	−0.079**	−0.082**	0.205**	0.029**	0.078**	0.083**	−0.076**
	(0.0025)	*(0.0050)*	*(0.0054)*	*(0.0216)*	*(0.0063)*	*(0.0195)*	*(0.0231)*	*(0.0384)*	*(0.0055)*	*(0.0158)*	*(0.0191)*	*(0.0313)*
Land	0.017**	0.016**	0.016**	−0.055**	0.021**	0.032**	0.033**	−0.130	−0.038**	−0.048**	−0.048**	0.185**
	(0.0009)	*(0.0011)*	*(0.0012)*	*(0.0199)*	*(0.0019)*	*(0.0033)*	*(0.0037)*	*(0.0945)*	*(0.0018)*	*(0.0031)*	*(0.0035)*	*(0.0797)*
Sample	A	A	A	A	A	A	A	A	A	A	A	A
Specification	(NULL)	(ALT1)	(ALT2)	(ALT3)	(NULL)	(ALT1)	(ALT2)	(ALT3)	(NULL)	(ALT1)	(ALT2)	(ALT3)
Year dummies	yes	yes			yes	yes			yes	yes		
Country effects		yes				yes				yes		
Cty-year dummies			yes	yes			yes	yes			yes	yes
Region effects				yes				yes				yes
Prob>F(NULL–ALT)	N/A	0.0000	0.0000	0.0000	N/A	0.0000	0.0000	0.0000	N/A	0.0000	0.0000	0.0000
Prob>F(ALL)	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
R-squared	0.40	0.63	0.65	0.96	0.29	0.40	0.42	0.97	0.45	0.53	0.55	0.98
Sum of coeff.	0.0003	−0.0046	−0.0050	−0.1730	0.0082	0.0239	0.0240	0.1180	−0.0085	−0.0192	−0.019	0.055
Linear homog.	*(0.7707)*	*(0.0000)*	*(0.0001)*	*(0.0000)*	*(0.0168)*	*(0.0001)*	*(0.0002)*	*(0.2763)*	*(0.0178)*	*(0.0005)*	*(0.0011)*	*(0.5373)*
(*p*-value)	accept	reject	reject	reject	reject	reject	reject	accept	reject	reject	reject	accept
Maddala–Wu	*(0.0188)*	*(0.0389)*	*(0.0263)*	*(0.0002)*	*(0.0040)*	*(0.0020)*	*(0.1949)*	*(0.1779)*	*(0.1705)*	*(0.2460)*	*(0.0464)*	*(0.4115)*
(*p*-value)	reject	reject	reject	reject	reject	reject	accept	accept	accept	accept	reject	accept

Note: 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. Standard errors in parentheses are heteroscedasticity-robust and clustered on country-year. ** denotes significance at the 5 per cent level, * denotes significance at the 10 per cent level.

in all industries, providing further evidence against the HO null hypothesis.

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, technology and other omitted variables does not substantially alter the relationship between factor endowments and regional production patterns, and that it is more important to control for time-invariant cross-country differences in relative prices, technology and other omitted variables in the move from (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 per cent level.

Column (4) of Table 2 (panels A–C) reports the results of extending the model 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, and any variation reflects either the reclamation of areas from the sea (the Netherlands) or measurement error, 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.

Table 2 also reports 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 in the order of magnitude of 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 panel data 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 misspecification. Two possible explanations for the non-stationarity of the residuals are, first, the omission of information on relevant factor endowments and, second, region-time variation in relative prices, technology and other omitted variables that have not been controlled for (and which will therefore be included in the error term).

5.2 Aggregate Industries, Disaggregate Endowments

Table 3 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).[12] In the interest 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 alternative, and we begin by considering the estimation results for specification (ALT2).

The estimated coefficients on physical capital are very similar to those 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

[12] The model in Table 2 (panels A–C) was re-estimated for the reduced sample; this yields very similar results to those reported in the paper.

Table 3: *Factor Endowments and Specialization at the Aggregate Level*

	(1) 696 obs. 1975–95	(2) 696 obs. 1975–95	(3) 696 obs. 1975–95	(4) 696 obs. 1975–95	(5) 696 obs. 1975–95	(6) 696 obs. 1975–95
Capital	−0.016**	0.083**	−0.067**	0.012**	0.059**	−0.071**
	(0.0051)	*(0.0244)*	*(0.0223)*	*(0.0057)*	*(0.0116)*	*(0.0114)*
Low education	0.028**	−0.035*	0.007	−0.040**	0.015	0.024
	(0.0074)	*(0.0215)*	*(0.0170)*	*(0.0097)*	*(0.0201)*	*(0.0207)*
Med. education	−0.030**	0.077**	−0.048**	−0.014**	0.021**	−0.007
	(0.0091)	*(0.0189)*	*(0.0142)*	*(0.0046)*	*(0.0097)*	*(0.0061)*
High education	−0.001	−0.130**	0.131**	−0.011**	0.015*	−0.004
	(0.0051)	*(0.0209)*	*(0.0189)*	*(0.0035)*	*(0.0086)*	*(0.0075)*
Arable land	0.012**	0.020**	−0.032**	−0.0010	0.017	−0.016
	(0.0011)	*(0.0032)*	*(0.0028)*	*(0.0019)*	*(0.0145)*	*(0.0149)*
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)*	*(0.0046)*	*(0.0959)*	*(0.0000)*	*(0.0020)*	*(0.0609)*
(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

Note: 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. Standard errors in parentheses are heteroscedasticity-robust and clustered on country-year. ** denotes significance at the 5 per cent level, * denotes significance at the 10 per cent level.

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 2B to 0.50 in column (2) of Table 3. Furthermore, we are now able to reject the null hypothesis of a unit root in the residuals at the 5 per cent 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 per cent 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 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.

Table 4 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 4 is similar to that reported for (ALT2) in Table 3. 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.[13]

The constancy of the estimated parameters as one moves from (ALT1) to (ALT2) in Table 2, the inclusion of country-year dummies in (ALT2)

[13] We also experimented with differencing over six- and eight-year time periods which yielded a broadly similar pattern of results.

Table 4: *Factor Endowments and Specialization at the Aggregate Level (Long Differences)*

	(1) 341 obs. 1975–95	(2) 341 obs. 1975–95	(3) 341 obs. 1975–95
Δ Capital	−0.006	0.062**	−0.057**
	(0.0049)	*(0.0120)*	*(0.0112)*
Δ Low education	−0.035**	−0.009	0.044**
	(0.0048)	*(0.0165)*	*(0.0195)*
Δ Med. education	−0.031**	0.023**	0.008
	(0.0058)	*(0.0078)*	*(0.0078)*
Δ High education	−0.010**	−0.002	0.012**
	(0.0020)	*(0.0042)*	*(0.0055)*
Δ Arable land	0.013**	0.018	−0.031**
	(0.0036)	*(0.0123)*	*(0.0132)*
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)*	*(0.0025)*	*(0.4462)*
(*p*-value)	reject	reject	accept

Note: 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. Standard errors in parentheses are heteroscedasticity-robust and clustered on country-year. ** denotes significance at the 5 per cent level, * denotes significance at the 10 per cent level.

which means the identification of factor endowment coefficients comes from 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 of the paper, we concentrate on the results using the more disaggregated data on factor endowments that controls for educational attainment and land quality.

The analysis so far has established that the strict version of the HO model with identical relative prices and technology is rejected against more general alternatives, although factor endowments are statistically significant and quantitatively important determinants of production structure, with a pattern of estimated coefficients broadly consistent with economic priors.

5.3 Disaggregate Industries, Disaggregate Endowments

In this section, we examine whether factor endowments are more or less successful in explaining specialization in disaggregated industries within the manufacturing sector than in the aggregate industries we have considered above. It is often argued in the literature on imperfect competition and trade (e.g. Helpman and Krugman 1985) that other considerations such as imperfect competition and increasing returns to scale are more important for production and trade within Manufacturing.

Table 5 reports the results of estimating our preferred specification (ALT2) with the five factor endowments for 11 disaggregated manufacturing industries, including Fuels, Ferrous Metals, Non-metallic Minerals, Chemicals, Machinery, Transport Equipment, Food Beverages & Tobacco, Textiles, Paper & Printing, Building & Construction, and Miscellaneous Other Production.

Factor endowments are again found to play a statistically significant role in explaining production structure, and the general equilibrium effects of factor endowments on production in individual industries vary substantially within manufacturing. Physical capital is positively and statistically significantly related to the share of Chemicals and Machinery in a region's GDP. Medium-education is positively and statistically significantly correlated with specialization in Metals, Machinery and Transport Equipment.

In all industries, the HO specification (NULL) is rejected against the more general alternative (ALT2) at the 5 per cent level of statistical significance, as shown in the first of the F-statistics reported towards the bottom of each column and panel of the table. The contributions of the country-year dummies are again not only statistically significant but also quantitatively important, with the regression R^2 in each industry rising markedly between specifications (NULL) and (ALT2).

In Table 6 we examine the predictive power of the model at different levels of industrial disaggregation by calculating within-sample average absolute prediction errors. These are defined in proportional terms as $\left|s_{zjt} - \hat{s}_{zjt}\right| / s_{zjt}$, where a hat above a variable indicates a predicted value.[14]

For the three aggregate industries of Agriculture, Manufacturing and Services, we find that the HO specification (NULL) provides a relatively

[14] 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).

Table 5: *Factor Endowments and Specialization at the Disaggregate Level (11 Industries)*

	(1) 696 obs. 1975-95	(2) 689 obs. 1975-95	(3) 689 obs. 1975-95	(4) 696 obs. 1975-95	(5) 696 obs. 1975-95	(6) 693 obs. 1975-95	(7) 696 obs. 1975-95	(8) 696 obs. 1975-95	(9) 696 obs. 1975-95	(10) 696 obs. 1975-95	(11) 696 obs. 1975-95
Capital	−0.015	−0.014**	−0.003**	0.007**	0.057**	0.005	0.012**	0.022**	0.015**	0.021**	−0.022**
	(0.0096)	(0.0027)	(0.0014)	(0.0015)	(0.0071)	(0.0030)	(0.0018)	(0.0038)	(0.0012)	(0.0024)	(0.0030)
Low education	−0.001	0.011**	0.002*	0.003	−0.051**	−0.008**	0.012**	0.003	−0.018**	−0.014**	0.027**
	(0.0132)	(0.0027)	(0.0013)	(0.0020)	(0.0123)	(0.0026)	(0.0040)	(0.0033)	(0.0023)	(0.0038)	(0.0054)
Med. education	−0.013	0.018**	0.005**	−0.001	0.049**	0.022**	−0.022**	0.018**	0.011**	0.017**	−0.031**
	(0.0153)	(0.0038)	(0.0017)	(0.0032)	(0.0194)	(0.0039)	(0.0054)	(0.0067)	(0.0032)	(0.0053)	(0.0062)
High education	0.002	−0.013**	−0.008**	−0.002	−0.041**	−0.012**	−0.011**	−0.035**	−0.005**	−0.021**	0.016**
	(0.0072)	(0.0024)	(0.0012)	(0.0024)	(0.0105)	(0.0028)	(0.0032)	(0.0054)	(0.0018)	(0.0032)	(0.0033)
Arable land	0.007**	0.001	0.002**	−0.001	0.002**	−0.001	0.004**	0.001	−0.001*	0.003**	0.003**
	(0.0023)	(0.0004)	(0.0003)	(0.0005)	(0.0009)	(0.0005)	(0.0004)	(0.0005)	(0.0002)	(0.0003)	(0.0005)
Industry	Fuel	Metal	Mineral	Chem	Machine	Transp	Food	Textile	Paper	Other	Constr
Sample	B	B	B	B	B	B	B	B	B	B	B
Specification	(ALT2)	(ALT2)	(ALT2)	(ALT2)	(ALT2)	(ALT2)	(ALT2)	(ALT2)	(ALT2)	(ALT2)	(ALT2)
Cty-year dummies	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Prob>F(NULL–ALT2)	0.0162	0.0000	0.0000	0.0000	0.0000	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	0.0000	0.0000	0.0000	0.0000	0.0000
R-squared	0.27	0.68	0.50	0.33	0.39	0.36	0.57	0.43	0.64	0.54	0.69
Sum of coeff.	−0.0191	0.0030	−0.0020	0.0060	0.0160	0.0059	−0.0051	0.0085	0.0027	0.0059	−0.0074
Linear homog.	(0.0000)	(0.0047)	(0.0158)	(0.0000)	(0.0000)	(0.0000)	(0.0002)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
(*p*-value)	reject	reject	reject	reject	reject	reject	reject	reject	reject	reject	reject
Maddala–Wu	(0.0101)	(0.0000)	(0.0000)	(0.1195)	(0.0538)	(0.0000)	(0.2104)	(0.0068)	(0.0017)	(0.0028)	(0.1405)
(*p*-value)	reject	reject	reject	accept	reject	reject	accept	reject	reject	reject	accept

Note: 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. Standard errors in parentheses heteroscedasticity-robust and clustered on country-year. ** denotes significance at the 5 per cent level, * denotes significance at the 10 per cent level.

Table 6: *Average Shares of Sectors in GDP and Within-Sample Average Absolute Prediction Errors*

Industry	(1) GDP share	(2) Prediction error (NULL)	(3) Prediction error (ALT2)
Aggregate industries			
Agriculture	0.042	0.675	0.582
Manufacturing	0.354	0.165	0.133
Services	0.604	0.088	0.068
Mean	0.333	0.309	0.261
Disaggregate industries			
Fuel	0.053	0.619	0.569
Metal	0.015	2.894	1.677
Mineral	0.016	0.308	0.295
Chemical	0.021	0.507	0.406
Machine	0.062	0.581	0.447
Transport	0.021	0.612	0.528
Food	0.040	0.355	0.255
Textile	0.022	1.040	0.811
Paper	0.015	0.415	0.288
Other	0.021	0.407	0.309
Construction	0.069	0.152	0.108
Mean	0.032	0.717	0.518
Sample	B	B	B

Note: Table reports mean values for the whole sample. Absolute proportional prediction errors are calculated as $|s - s(P)|/s$, where a capital P indicates a predicted value. Prediction error (NULL) is based on the fitted values from specification (NULL) using the disaggregated data on five factor endowments and Sample B. Prediction error (ALT2) is based on the fitted values from specification (ALT2) using the disaggregated data on five factor endowments and Sample B. Reported prediction errors exclude region Be1 (Brussels). Brussels is a capital city and the shares of some industries in this region are clear outliers. As a robustness test, we reestimated the model excluding this region. This produced very similar estimated coefficients to those reported earlier. Reported prediction errors also exclude the Metal industry in region Esp7 (Canaries). This industry constitutes a very small share of GDP in this region. As a robustness test, we reestimated the model excluding this region. This produced very similar estimated coefficients to those reported earlier.

successful explanation for regional patterns of production. The average absolute prediction error in Manufacturing of 16.5 per cent compares favourably with values reported using country-level data. The model is most successful in explaining regional production patterns in Services and least successful in Agriculture, which is consistent with time-varying policy interventions that have uneven effects across European regions being important in Agriculture. The introduction of the country-year dummies in our preferred specification (ALT2) leads to a reduction in the model's within-sample average absolute prediction error, with the greatest impact observed in Agriculture.

Both the HO specification (NULL) and the alternative specification (ALT2) are less successful in explaining regional patterns of production in disaggregated industries within Manufacturing. Across the disaggregated manufacturing industries as a whole, the average absolute prediction error rises to 71.7 per cent in specification (NULL) and 51.8 per cent in specification (ALT2). The model is least successful in explaining patterns of production in Metal Products (which includes Iron & Steel) and Textiles, both of which are again industries where historically there has been extensive policy intervention in Europe.

These findings of greater within-sample average absolute prediction errors in disaggregated manufacturing industries are consistent with the larger number of industries relative to factors of production at the disaggregated level, suggesting the possibility of production indeterminacy. They are also consistent with the idea that omitted regional variation in relative prices, technology or other determinants of production structure may be particularly large in individual disaggregated manufacturing industries.

6 Conclusions

This paper has analysed the relationship between production structure and factor endowments using panel data on 14 industries in 45 regions from 7 European countries since 1975. We derived a general equilibrium relationship linking the share of a sector in GDP to factor endowments, relative prices and technology from the neoclassical model of trade. The HO model is a special case of this framework where relative prices and technology are identical across regions. We compared the empirical performance of the HO null against a series of more general alternatives allowing for variation in relative prices, technology and other variables omitted from the HO model.

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, measurement error and technology differences are likely to be much smaller across regions within Europe than for a broad cross-section of developed and developing countries. The data also have the advantage of including both manufacturing and non-manufacturing sectors, where the latter play an important role in European regions, and we are able to explore the model's empirical performance for both broad and more finely detailed industry classifications.

For both the three aggregate industries (Agriculture, Manufacturing and Services) and the 11 disaggregated industries within Manufacturing, the strict version of the HO model assuming identical relative prices and technology is rejected against more general alternatives. Nonetheless, factor endowments play a statistically significant and quantitatively important role in explaining production structure in European regions, with a pattern of estimated coefficients broadly consistent with economic priors.

The introduction of more disaggregated factor endowment measures that control for educational attainment and land quality leads to a substantial improvement in the model's explanatory power. The HO null specification is more successful in explaining the production structure for broad industries, such as Agriculture, Manufacturing and Services, than for more narrowly defined industries within manufacturing. This is consistent with a larger number of industries relative to factors of production at the disaggregate level, introducing the possibility of production indeterminacy, and with the idea that omitted regional variation in relative prices, technology or other determinants of production structure may be particularly large in individual manufacturing industries.

Appendix A

Table A1: *Sample Composition*

Country	Sample A	Sample B	Number of NUTS-1 regions
Belgium	1975–1995	1979–1995	3 (be1–be3)
Spain	1980–1995	1980–1994	7 (esp1–esp7)
France	1975–1995	1977–1994	8 (fra1–fra8)
Italy	1975–1995	1980–1995	11 (ita1–ita9, itaa/b)
Luxembourg	1975–1995	1979–1990	1 (lux)
Netherlands	1975–1995	1977–1995	4 (ndl1–ndl4)
United Kingdom	1975–1986	1975–1986	11 (uk1–uk9, uka/b)

Table A2: *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 A3: *Regions Included in the Sample*

Code	Description	Code	Description
	Belgium		*Italy*
Be1	Brussels	Ita1	North West (I)
Be2	Vlaams Gewest	Ita2	Lombardia
Be3	Region Wallonne	Ita3	Nord East (I)
		Ita4	Emilia-Romagna
	Spain	Ita5	Centre (I)
Esp1	North West (E)	Ita6	Lazio
Esp2	North East (E)	Ita7	Abruzzo-Molise
Esp3	Madrid	Ita8	Campania
Esp4	Centre (E)	Ita9	South (I)
Esp5	East (E)	Itaa	Sicily
Esp6	South (E)	Itab	Sardinia
Esp7	Canaries		
			Luxembourg
	France	Lux	Luxembourg (Grand-Duche)
Fra1	Ile De France		
Fra2	Bassin Parisien		*United Kingdom*
Fra3	Nord-Pas-de-Calais	UK1	North (UK)
Fra4	East (F)	UK2	Yorkshire and Humberside
Fra5	West (F)	UK3	East Midlands
Fra6	South West (F)	UK4	East Anglia
Fra7	Centre-East (F)	UK5	South East (UK)
Fra8	Mediterranean	UK6	South West (UK)
		UK7	West Midlands
	Netherlands	UK8	North West (UK)
Nld1	North-Netherland	UK9	Wales
Nld2	East-Netherland	UKA	Scotland
Nld3	West-Netherland	UKB	Northern Ireland
Nld4	South-Netherland		

Appendix B

B1. Region-Level Data on Production and Endowments

1. Value Added: current price value-added, millions of ECUs, from Regio data set, Eurostat.
2. GDP: current price, millions of ECUs, from Regio data set, Eurostat.
3. Population: total population, thousands of people, from Regio data set, Eurostat.
4. Land: total land area, thousands of hectares, from Regio data set, Eurostat.
5. Arable Land: total arable land area, thousands of hectares, from Regio data set, Eurostat.
6. Capital Stock: constructed by the perpetual inventory method using region-level investment data (Gross Fixed Capital Formation), constant 1990 prices, millions of ECUs. The main source for the investment data is the Regio data set, 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, assuming that the growth in investment in each region is the same as the growth in investment in the country as a whole for the years where regional data are not available.

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 national labour force surveys, we compute the percentage of the population in each region with each level of educational attainment. The endowment variables included in the regressions are these percentages multiplied by the regional 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 1983–1994. Linear interpolation of the data when required.
3. Italy: educational attainment data from 1986–1997 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 the country-level data in Nickell et al. (2001), assuming that the growth in educational attainment in each region is the same as the growth in educational attainment in the country as a whole for the years where regional data are not available.

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 using the country-level data in Nickell et al. (2001), assuming that the growth in educational attainment in each region is the same as the growth in educational attainment in the country as a whole for the years where regional data are not available.
5. Netherlands: Data from *National Statistical Office*, years 1992–1998. The regional data are extended backwards in time using country-level information from Nickell et al. (2001), assuming that the growth in educational attainment in each region is the same as the growth in educational attainment in the country as a whole for the years where regional data are not available.
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 1983–1994. 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.

References

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

Bowen, H., E. Leamer, and L. Sveikauskas (1987). Multicountry, Multifactor Tests of the Factor Abundance Theory. *American Economic Review* 77 (5): 791–809.

Cameron, G., and J. Muellbauer (1998). The Housing Market and Regional Commuting and Migration Costs. *Scottish Journal of Political Economy* 45 (4): 420–446.

Davis, D., D. Weinstein, S. Bradford, and K. Shimpo (1997). Using International and Japanese Regional Data to Determine When the Factor Abundance Theory of Trade Works. *American Economic Review* 87 (3): 421–446.

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

Gabaix, X. (2005). The Factor Content of Trade: A Rejection of the Heckscher–Ohlin–Leontief Hypothesis. Mimeo, MIT.

Gandal, N., G. Hanson, and M. Slaughter (2004). Technology, Trade, and Adjustment to Immigration in Israel. *European Economic Review* 48 (2): 403–428.

Griliches, Z., and J. Hausman (1986). Errors in Variables in Panel Data. *Journal of Econometrics* 31 (1): 93–118.

Hanson, H., and M. Slaughter (2002). Labor Market Adjustment in Open Economies: Evidence from US States. *Journal of International Economics* 57 (1): 3–29.

Harrigan, J. (1995). Factor Endowments and the International Location of Production: Econometric Evidence for the OECD, 1970–1985. *Journal of International Economics* 39 (1–2): 123–141.

Harrigan, J. (1997). Technology, Factor Supplies, and International Specialization: Estimating the Neoclassical Model. *American Economic Review* 87 (4): 475–494.

Helpman, E. (1984). Increasing Returns, Imperfect Markets, and Trade Theory. In R. Jones and P. Kenen (eds.), *Handbook of International Economics.* Volume I. Amsterdam: North-Holland.

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

Im, K., M. Pesaran, and Y. Shin (2003). Testing for Unit Roots in Heterogeneous Panels. *Journal of Econometrics* 115 (1): 53–74.

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

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

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

Machin, S., and J. Van Reenen (1998). Technology and Changes in Skill Structure: Evidence from Seven OECD Countries. *Quarterly Journal of Economics* 113 (4): 1215–1244.

Maddala, G., and S. Wu (1999). A Comparative Study of Unit Root Tests with Panel Data and a New Simple Test. *Oxford Bulletin of Economics and Statistics.* Special Issue: 631–652.

McCormick, B. (1997). Regional Unemployment and Labour Mobility in the United Kingdom. *European Economic Review* 41 (3–5): 581–589.

Nickell, S., and B. Bell (1996). Changes in the Distribution of Wages and Unemployment in OECD Countries. *American Economic Review* 86 (2): 302–308.

Nickell, S., S. Redding, and J. Swaffield (2001). Educational Attainment, Labour Market Institutions and the Structure of Production. CEPR Discussion Paper 3068. Centre for Economic Policy Research, London.

Pedroni, P. (1999). Critical Values for Cointegration Tests in Heterogeneous Panels with Multiple Regressors. *Oxford Bulletin of Economics and Statistics.* Special Issue: 653–670.

Pesaran, M., Y. Shin, and R. Smith (1999). Pooled Mean Group Estimation of Dynamic Heterogeneous Panels. *Journal of the American Statistical Association* 94 (446): 621–634.

Quah, D. (1994). Exploiting Cross-Section Variation for Unit Root Inference in Dynamic Data. *Economics Letters* 44 (1–2): 9–19.

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

Schott, P. (2003). One Size Fits All? Heckscher–Ohlin Specialization in Global Production. *American Economic Review* 93 (2): 686–708.

Trefler, D. (1995). The Case of the Missing Trade and Other Mysteries. *American Economic Review* 85 (5): 1029–1046.