

Urban versus Rural Return to Human Capital in Portugal: a cook-book recipe for applying assignment models

C. N. Teulings *

J. A. C. Vieira **

Abstract

The Portuguese economy has been characterised by modernisation since the post-war period. Lisbon and the Tagus Valley is a centre of this process. Hence, this region faces a high demand for highly skilled labour. This paper analyses rates of return on human capital in the region of Lisbon and in the rest of the country in the period 1982-1992. An assignment model of heterogeneous workers to heterogeneous jobs is discussed. We also develop a cook-book recipe for its estimation. The main implication, a high return being associated with similar workers being assigned to more complex jobs, is born out by the data. Apparently new technology in Lisbon triggers the demand for human capital, not the other way around. The estimation procedure allows us to obtain a simple, free dimension, parameter that measures the heterogeneity in jobs and therefore the ease of substitution between worker types. This is called the complexity dispersion parameter. Calculations suggest that paying half the optimal wage level is about to double the cost per efficiency unit of labour.

* University of Amsterdam and Tinbergen Institute
address: Keizersgracht 482, 1017 EG Amsterdam
e.mail: teulings@few.eur.nl

** University of the Azores and Tinbergen Institute
address: University of the Azores, Dep. of Economics, 9500 P. Delgada, Portugal
e.mail: jvieira@alf.uac.pt

The authors thank Joop Hartog for his helpful comments.

1 Introduction

The Portuguese economy has been subject to deep transformations during the last fifty years. After the second world war it developed from a rural to a service economy. In 1960, the economy relied heavily on agriculture and fishery, which were responsible for 42.8% of total employment. Manufacturing and services contributed 29.5% and 27.7%, respectively. By the time of the revolution in 1974 the numbers were 34.9% for agriculture and fishery, 33.8% for manufacturing, and 31.3% for services. From then on, modernization speeded up. By 1991, agriculture and fishery accounted for only 10.8% of employment, while manufacturing grew slightly to 37.9% and the share of services expanded to 51.3%. At the same time, the population has become more educated. Illiteracy rates within the working population aged 15-64 years decreased from 33.9% in 1960 to 6.5% in 1991.

In the initial phase after World War II, economic activities were protected through a deliberate licensing policy called *industrial conditioning*. As time went by, this policy promoted the emergence of a dualistic industrial structure with a large number of small firms intensive in unskilled labour coexisting with a few large conglomerates in modern industries (e.g. chemical products, banking and insurance). That duality is also mirrored in the distribution of activities across regions. The region of Lisbon and the Tagus Valley (hereafter LTV) captured most of the high-tech and financial activities and has become a centre of rapid transformation. In order to satisfy the quickly growing demand, labour flew from the country side to the capital. In 1991, 33.4% of the population lived in LTV as compared with 28.7% in 1970 or 22.6% in 1940. Approximately 34.5% of the employment was in or around Lisbon in 1991. Of this, 5.1% were employed in agriculture and fishing, 30.2% in manufacturing, and 64.7% in services. For the rest of the country, the distribution was 13.8% in agriculture and fishing,

41.9% in manufacturing and 44.3% in services. Hence, LTV took the lead in particular in services.

Clearly, the modernisation in LTV affected not only the size but also the composition of the workforce. Where new technologies requiring skilled labour were concentrated in LTV, labour demand was skewed towards highly skilled workers (see some figures for 1982 and 1992 in Table 1). What force drives this process? Does the supply of new better educated generations sparks off the introduction of new technologies, or is the process demand driven and does the new high-tech economy pull the highly skilled workers to the LTV region? The analysis of this modernisation process can provide information on more structural features of the degree of substitutability in the skill composition of labour demand.

The model we apply for the analysis is a simplified version of the assignment model discussed in Teulings (1995a). That model considers the assignment of heterogeneous workers to heterogeneous jobs. There are both an infinite number of worker and job types. Worker types are characterised by a single index, referred to as the skill level of the worker, which varies continuously along the real axis. Worker's skill level will be measured by years of education and experience. Likewise, job types are characterised by an index of complexity, which can take any positive number. Dummies for occupation and industry will be used to measure job complexity. The assignment of workers to jobs and relative wages for various worker types are analysed in a competitive market equilibrium, where firms pick the optimal worker type for their jobs depending on the relative wages for each skill level.

Though the estimation methodology for this model proposed in Teulings (1995a) is fairly general, it suffers the disadvantage of being non-linear and therefore complicated and not very transparent. In this paper, we offer simple alternative which just requires four OLS regressions. Furthermore, we propose a simple, dimension free parameter, that measures the heterogeneity

in jobs and therefore the ease of substitution between different worker types, that can be compared between economies either in time or between countries. We shall refer to this parameter as the complexity dispersion parameter. When complexity dispersion is low and all workers occupy therefore more or less the same job, worker types can be substituted for each other easily, without much affecting relative wages. When complexity dispersion is high, worker types will be less easily substitutable. The complexity dispersion parameter has an appealing economic interpretation. It measures the additional cost per efficiency unit of labour when a firm hires not exactly the optimal worker type for a particular job. The higher the dispersion of job complexity, the more costly are deviations from the optimal allocation. We shall compare the estimate of the complexity dispersion parameter for Portugal with estimates that can be derived from previous studies on the Netherlands and the United States.

The paper is organised as follows. Section 2 sets out the basics of the theoretical model. The estimation methodology is discussed in section 3. Section 4 deals with the estimation results. Finally, section 5 concludes and summarises.

2 The model

In an economy t , there is an infinite number of worker and job types. Workers are assigned to job types and vice versa on the labour market. The distribution of worker types among labour supply is determined exogenously. Each job type produces its own commodity, traded on commodity markets. For simplicity, we assume that each firm offers jobs of only a single type. The latter assumption is merely a matter of presentation and can be discarded without changing the implications of the model. For the purpose of this paper, the precise formulation of commodity demand is immaterial to the results. Finally, there is perfect competition on all markets.

The labour market model relies on four specific assumptions:

- 1) a single index for human capital: all worker characteristics can be summarized in a single index referred to as the level of skill s , $s \in (-\infty, \infty)$;
- 2) idem for jobs: all job characteristics can be summarized in a single index referred to as the job complexity c , $c \in (0, \infty)$;
- 3) absolute advantage: within a job type, better skilled workers are always more productive;
- 4) comparative advantage: the relative productivity gain of additional units of skill goes up with the level of job complexity.

More specifically, assumption 3) and 4) can be captured by the following specification of the productivity $\Pi(c,s)$ of worker type s in job type c :

$$\Pi(c,s) = \exp[c s]. \tag{1}$$

Absolute advantage is implied by: $d \Pi(\cdot)/ds = c \Pi(\cdot) > 0$, due to the assumption that c is strictly positive. Comparative advantage is implied by $d\{ d \ln \Pi(\cdot)/ds \}/dc = 1 > 0$. Let $w_t(s)$ be the log wage for worker type s in the market equilibrium for economy t . The problem of a firm offering jobs of type c is to minimize the cost per efficiency unit of labour by hiring the right worker type $s_t(c)$:

$$s_t(c) = \operatorname{argmin}_s \exp[w_t(s) - c s].$$

Note that in the absence of obvious units of measurement, the model is unaffected by any linear transform of s , e.g. $s^* = \mathbf{S} s$, provided that an offsetting transform of c is applied, i.e. $c^* = \mathbf{S}^{-1} c$.¹

The first order condition defines the optimal skill type as a function of the complexity of the jobs offered by the firm:

$$w_t'[s_t(c)] = c, \quad \text{for all } c. \tag{2}$$

The first order condition implies $w_t'(\cdot)$ to be strictly positive because $c > 0$ by assumption. This is the implication of assumption 3), absolute advantage.

The second order condition is:

$$w_t''[s_t(c)] > 0, \quad \text{for all } c. \tag{3}$$

----- insert Fig. 1 about here -----

The situation is depicted in Figure 1. The line $c s$ in the upper half is the equal-profit line (per unit of output). The optimum is reflected by the point of tangency of the locus $w_t(s)$ and this equal-profit

¹The model remains also unchanged by the more general transform: $s^* = \mathbf{S}_0 + \mathbf{S}_1 s$. The effect of the intercept can be undone by changing the units of measurement of product type c by $-\mathbf{S}_0 c$.

line. The lower half depicts the marginal conditions. $w_i'(\cdot)$ is upward sloping when the second order condition is satisfied. Assume to the contrary that $w_i''(\cdot)$ is negative for some subset (s_1, s_2) of the domain of s , see Figure 2. An even larger subset (s_1, s_{2+}) of the domain will not be employed by any firm type c in that case. When there is a positive supply of these types, market forces will drive down wages. Hence, the situation depicted in Figure 2 can never be an equilibrium outcome of the model. When $w_i''(\cdot)$ is equal to zero for some subset, all worker types in this subset will be employed in the same job type, for which $w_i'(\cdot) = c$. This can only be true when there is mass point in the distribution of product demand among c -types. Assuming a distribution without mass points rules out this case.

----- insert Fig. 2 about here -----

Differentiating the first order condition (2) with respect to c yields:

$$w_i''(\cdot) s_i'(\cdot) = 1.$$

Hence, $s_i'(\cdot)$ is strictly positive by the second order condition and the function $s_i(\cdot)$ therefore has a well defined inverse function, denoted $c_i(\cdot)$. Condition (2) can then be written as:

$$w_i'(s) = c_i(s), \quad \text{for all } s. \tag{4a}$$

and hence, from the second order condition:

$$w_i''(s) = c_i'(s) > 0, \quad \text{for all } s. \tag{4b}$$

The first order condition has a clear economic interpretation. $w_t'(s)$ measures the return on human capital in economy t . Each additional unit of s yields $w_t'(s) \times 100\%$ higher wage. When the supply of highly skilled workers is limited in economy $t=1$ compared to economy $t=2$, then jobs of complexity c will be assigned to less skilled workers in economy 1 than in economy 2, simply because the better skilled workers are scarce in economy 1. At the same time, the return to human capital will be higher in economy 1. This is exactly what condition (4a) states. There is a one-to-one relation between the return on human capital and the allocation of skills to job complexities.

Equation (4b) has three important economic implications. First, better skilled workers are assigned to more complex jobs in market equilibrium. $w_t''(\cdot)$ being positive is a direct implication of comparative advantage, a conclusion that has been first noticed by Sattinger (1975). The distribution of wages is therefore always skewed to the right compared to distribution of productivity within any job type. A better skilled worker gets higher wages than his less skilled colleague, first because he is more productive when he works in the same job type, and second because he is assigned to a more complex job. Transferring his less skilled colleague to the more complex job would be counter productive, because the skill differential with the highly skilled worker would have a larger negative effect on productivity there.

Secondly, the equality $w_t''(\cdot) = c_t'(\cdot)$ implies the second derivative of the wage function to be a measure of the dispersion of job complexity c among demand (taking the dispersion of s as fixed). When $w_t''(\cdot)$ would be zero for the full domain of s , all jobs would be of equal complexity and types of labour would be perfectly substitutable. A rise in wages for some skill types will have large effects on their employment because all workers are employed in the same job type and productivity differentials are fixed within a job type; then: $w_t(s) = \mathbf{q} + c_0 s$, where \mathbf{q} is some constant and c_0 is the complexity of the unique job type. The larger $w_t''(\cdot)$, the smaller will be the effect of changes in

relative wages on the assignment of workers to jobs and the smaller will therefore be the elasticities of substitution.

Figure 3 provides an intuition by depicting two polar cases. The line cs is a zero profit line for product type c . Now, suppose that demand for this commodity goes up, such that its price rises by Δ . The new zero profit line will then be $cs + D$. All worker types between both intersections of $w(s)$ and this new zero profit line can be profitably employed for the production of commodity c . Hence the more curved $w(s)$ the smaller the rise in output of commodity c for a given price increase, Δ , and the smaller is therefore the substitutability of worker types.

----- insert Fig. 3 about here -----

Finally, the second derivative is a measure for the additional cost per efficiency unit of labour due to an out-of-equilibrium assignment of workers to a particular firm type. A second order Taylor expansion around the optimum $s_i(c)$ of additional unit cost as a share of the efficient unit cost level yields: $\exp[w_i(s)-c s] / \exp\{w_i[s_i(c)]-c s_i(c)\} - 1 = \frac{1}{2} w_i''[s_i(c)] [s-s_i(c)]^2$. Therefore, hiring a worker type that earns k % more or k % less than the optimal worker type² raises cost by:

$$\frac{1}{2} \{ w_i''(.) / w_i'(.)^2 \} \times k^2 \cdot 100^{-1} \%$$

We shall refer to the statistic in curly brackets as the *complexity dispersion parameter*. Clearly, since the above formula depends on a second order Taylor expansion, it applies only locally. When we assume the complexity dispersion parameter to be constant along the domain of s (as we will do in the empirical application below), then it is easy to check that $w(s)$ satisfies:

$$w(s) = -\frac{1}{\mathbf{g}} \ln(s^+ - s) + \mathbf{y} \quad (5)$$

where γ is the complexity dispersion parameter. This expression can be used to calculate an exact formula instead of a Taylor expansion of the cost of paying a wage $\mathbf{D}w$ above the optimal wage $w(s)$. By paying a $\mathbf{D}w$ higher wage, the firm is able to hire a better skilled and hence more productive worker. Let $\mathbf{D}s$ be the skill differential. The additional productivity is $c(s)\mathbf{D}s = w'(s)\mathbf{D}s$. By equation

$$(5) \text{ we have: } \mathbf{D}w = -\frac{1}{\mathbf{g}} \ln(s^+ - s - \mathbf{D}s) + \frac{1}{\mathbf{g}} \ln(s^+ - s). \text{ Therefore: } \mathbf{D}s = [1 - \exp(-\mathbf{g}\mathbf{D}w)](s^+ - s).$$

Since: $w'(s) = \frac{1}{\mathbf{g}}(s^+ - s)^{-1}$, a general formula for the relative cost disadvantage is:

$$\mathbf{D}w - w'(s)\mathbf{D}s = \mathbf{D}w - \frac{1}{\mathbf{g}}[1 - \exp(-\mathbf{g}\mathbf{D}w)] \quad (6)$$

This formula will be applied in section 4.

Note that, arbitrary measurement of s is inconsequential, as this unit does not affect the complexity dispersion parameter. When applying the linear transform of s to s^* considered before, the parameter \mathbf{s} drops out in this case. The higher the complexity dispersion parameter, the more dispersed is job complexity, the less substitutable are skill types and the higher are the additional cost of out-of-equilibrium wage setting. Estimating this parameter is a central goal of this paper.

²Hence: $|s - s_i(c)| w_i'[s_i(c)] = k \%$.

3 Estimation methodology

In practice, we have no direct observations on s and c , so we shall have to infer them from observables, for example: $s_i = \underline{S}(x_i)$, where x_i is a vector of (partially unobserved) worker characteristics of worker/job combination i and where s_i is the skill level of the worker, and mutatis mutandis the same for the job characteristics z_i : $c_i = \underline{C}(z_i)$. Each worker/job combination i is observed only in a single economy t . We could therefore add a suffix t to denote the economy where i is observed, and will do so subsequently whenever convenient. Typically, x_i contains the variables showing up in ordinary Mincerian earnings equation: years of education and experience, sex and the like, and an unobserved component. The vector z_i contains variables like dummies for occupation and industry, log firm size, and again an unobserved component.

The functions $S(\cdot)$ and $C(\cdot)$ are taken identical for different economies t . This assumption is critical. It provides the empirical content of the single index assumptions 1) and 2) in section 2: worker and job characteristics can each be meaningfully aggregated in a single index. The assignment of workers to jobs, $c_i(s)$, varies between t due to shifts in either the distribution of s among labour supply or c among labour demand, and so does therefore the wage function. The relation between characteristics on the one hand and s and c on the other hand is, however, part of production technology and therefore exogenous to the model.

We provide a cook-book recipe that allows us to use four simple OLS regressions to estimate this model. By previous definitions, we have:

$$w_i(s_i) = w_i[\underline{S}(x_i)] \equiv \underline{W}_i(x_i), \tag{7}$$

We cannot simply assume as a first order approximation that both $\underline{S}(\cdot)$ and $w_t(\cdot)$ are linear because we want to estimate the second derivative of $w_t(\cdot)$ while the simple linearity assumptions would impose a value of this second derivative. Moreover, the linearity assumption would be inconsistent with the previous discussion since it implies $w_t''(\cdot) = 0$, which contradicts the second order condition (3). However, there is no objection against assuming $\underline{W}_t(\cdot)$ to be linear, like we do routinely when estimating Mincerian earnings equations:

$$w_{it} = \mathbf{b}_{0t} + x_{it}'\mathbf{b}_t + \underline{x}_{it},$$

where \mathbf{b}_{0t} is a dummy for economy t , \mathbf{b}_t is a vector of parameters for economy t and where \underline{x}_{it} captures the unobserved characteristics of worker i . The non-linearity in $w_t(\cdot)$ will be picked up by a countervailing non-linearity in $\underline{S}(\cdot)$. The above equation is linear and can be estimated by OLS. However, the single index assumption implies: $\mathbf{b}_t = \mathbf{a}_t \mathbf{b}$, where \mathbf{b} is a vector of parameters and where \mathbf{a}_t is an economy variant scalar. This relation implies a multiplicative and hence non-linear restriction over the parameters. Clearly, one can deal with this problem by applying non-linear least squares techniques, but our cook-book recipe provides a simpler approach. First, we estimate:

$$w_{it} = \mathbf{b}_{0t} + x_{it}'\mathbf{b} + \underline{x}_{it} \tag{8a}$$

which can be estimated by OLS. Now define the human capital index: $q_{it} = x_{it}'\mathbf{b} + \underline{x}_{it}$ and its observed component $q_{it}^* = x_{it}'\mathbf{b}$. Then the equation:

$$w_{it} = \mathbf{a}_{0t} + q_{it}^*\mathbf{a}_{1t} + \underline{x}_{it} \tag{8b}$$

is estimated, which again can be done by OLS.

Define the functions: $S(q) \equiv \underline{S}(x)$ and $W_t(q) \equiv \underline{W}_t(x)$. Equation (8b) implies: $W_t'(\cdot) = \mathbf{a}_{1t}$. Hence, \mathbf{a}_{1t} measures the return on the human capital index q_{it} . Note that when $\mathbf{a}_{1t} = \mathbf{a}$ for all t , equation (8a) implies $W_t'(\cdot) = 1$. Therefore, we can define some properly weighted average of \mathbf{a}_{1t} and hence $W_t'(\cdot)$ to be equal to unity. The index q has therefore two interpretations. The classical interpretation is that it is Mincerian human capital, measuring the relative payoffs of the human capital components of years of education and years of experience. The new interpretation in the framework of the assignment model of this paper is that q is just a properly normalised non-linear transform of s , such

$$\text{that: } W_t''(q) = d^2\{w_t[S(q)]\}/(dq)^2 = w_t''[S(q)]S'(q)^2 + S''(q)w_t'[S(q)] = 0$$

$$\text{and hence: } S''(q) = -S'(q)^2 w_t''[S(q)]/w_t'[S(q)] < 0.$$

Therefore, estimating the non-linearity of $w_t(s)$ is equivalent to estimating $S''(q)$.

Substituting $S(q_i)$ and $W_t(q_i)$ for $\underline{S}(x_i)$ and $\underline{W}_t(x_i)$ in equation (5) and differentiating with respect to q yields:

$$W_t'(q) = w_t'[S(q)] S'(q). \tag{9}$$

A convenient specification of $S(\cdot)$ with flexibility in its second derivative is:

$$S(q) \equiv -^{1/g} \exp[-\mathbf{g}q - \gamma_1] + \mathbf{g}. \tag{10}$$

The parameter \mathbf{g} has a neat interpretation. Differentiating (9) with respect to q yields: $W_t''(q) = w_t''[S(q)] S'(q)^2 + w_t'[S(q)] S''(q)$. Using: $W_t'(q) = 1$ (on average across t) and: $W_t''(q) = 0$ and applying equation (9) and (10) yields: $w_t''[S(q)]/w_t'[S(q)]^2 = -S''(q)/S'(q) = \gamma$. Hence, \mathbf{g} is equal to

the complexity dispersion parameter in our linearised cook-book recipe version of the model. The second order condition (3) is equivalent to the restriction $\mathbf{g} > 0$.

Now, we apply the corner stone of our assignment model, first order condition (2), which specifies a relation between the assignment of worker types to jobs, $c_t(s)$, and the return to skill, $w_t'(s)$. Using the notation from this section, we have:

$$W_t'(q_{it}) = w_t'[S(q_{it})]S'(q_{it}) = \underline{C}(z_{it})S'(q_{it}).$$

Now, using $W_t'(q_{it}) = \mathbf{a}_{1t}$ and $S'(q) = \exp(-\gamma q - \gamma_1)$, and taking logs yields:

$$\ln \underline{C}(z_{it}) = \mathbf{g}_t + q_{it}\mathbf{g} + \ln \mathbf{a}_{1t}. \quad (11a)$$

Like $\underline{S}(x)$, $\ln \underline{C}(z)$ is assumed to be linear:

$$\ln \underline{C}(z_{it}) = \mathbf{d}^*(z_{it}'\mathbf{d} + z_{it}) \quad (11b)$$

where \mathbf{d} is a parameter vector and z_{it} are unobserved job characteristics. The reason for introducing the multiplicative parameter \mathbf{d}^* will become apparent later on.

Like for the supply side of the model, our cook-book recipe uses a two step procedure for the demand side. First, an estimate for \mathbf{d} is needed. By comparative advantage, we have a relation between the levels of complexity and of skill: $c_t'(s_{it}) = c_t'[S(q_{it})] > 0$. The latter specification is useful, since at this stage we are capable of calculating q_{it}^* , not s_{it} . Hence, the following equation is estimated for all t :

$$q_{it}^* = \mathbf{d}_{0t} + z_{it}'\mathbf{d} + \underline{z}_{it}^* \quad (12)$$

where the intercepts \mathbf{d}_{0t} have to account for the fact that function $c_t(\cdot)$ differs between economies t .³ The estimated value of \mathbf{d} is used to calculate: $j_{it} \equiv z_{it}'\mathbf{d}$. Again, the use of a single parameter vector \mathbf{d} to transform a vector of job characteristics in a scalar is motivated by the single index assumption for job characteristics. Equation (12) provides the weights for the aggregation of job characteristics in a single index, like equation (8a) does for worker characteristics. However, $\ln c_{it}$ can be any increasing function of the index j_{it} . Due to the linearization in our cook-book recipe, this function is restricted to the linear case with \mathbf{d}^* as the scaling factor between $\ln c_{it}$ and j_{it} , but that does not affect the basic argument.

Finally, the following equation is estimated:

$$j_{it} = \mathbf{q}_0 + q_{it}\mathbf{q}_1 + \ln \mathbf{a}_{1t}\mathbf{q}_2 + \underline{z}_{it} \quad (13)$$

where the estimated values of \mathbf{a}_{1t} from equation (6b) are applied.

A further analysis of the implications of (13) is revealing. From equations (11a), (11b) and (12), we have: $\mathbf{d}^* = 1/\mathbf{q}_2$, and: $\mathbf{g} = \mathbf{q}_1/\mathbf{q}_2$. Hence, equation (13) identifies the parameters of the transformation between $\ln c_{it}$ and j_{it} . Basically, the equation compares variations in the return to the human capital index \mathbf{a}_{1t} between economies t on the hand and in the human capital index q_{it} within an economy on the other hand. The higher the return to the human capital index in an economy, the more precious will be highly skilled workers and the lower will be the skill level of the workers hired

³In principle, instead of q_{it}^* one could also use w_{it} as the left hand variable in this equation because the model implies the latter variable to be positively related to c_{it} , too. However, when there also some elements of rents incorporated in

in a particular job type. When \mathbf{a}_{1t} rises by 10 % from $t=1$ to $t=2$ (e.g. $\mathbf{a}_{11}=1.0$ and $\mathbf{a}_{12}=1.1$), the level of q_{it} will be $10\%/g$ lower. Hence, the larger the variation in the assignment of worker type q_{it} to a particular job type relative to the variation in the return to human capital, the smaller is g . This squares well with the interpretation of γ as the complexity dispersion parameter and pictures in Figure 3. Note that g cannot be identified from data on only a single economy because \mathbf{q}_0 and \mathbf{q}_2 cannot be distinguished, see Kahn and Lang (1988). This conclusion applies also in the more general version of the model, see Teulings (1995a, p.301).

4 Estimation Results

The data used in the following estimation were drawn from Quadros Pessôal and cover the years of 1982, 1986 and 1992.⁴ These are administrative data: all firms with wage earners must complete a standardised questionnaire every year and send it to the Department of Labour. This is carried out in March of each year. The data include information on individual characteristics such as age, the highest level of schooling completed and gender. They also include information on variables such as occupation, industry affiliation, firm size, and job level. The latter variable ranks jobs on a seven point scale. Level 1 is called ‘very simple’, level 7 is called ‘scientific’. Information is also available for each worker on gross monthly earnings and hours worked per month. Civil servants and people serving the armed forces are not covered by the data. Missing values were deleted from the original samples as well as

wages in particular job types (e.g. a dummy for unionisation or, less clearly, firm size), then it is better to use q_{it}^* as the variable to explain this variable is less likely to be correlated to the rent share than wages itself.

⁴ The autonomous regions of the Azores and Madeira are not considered. Therefore, the data only refer to the Portuguese mainland.

agricultural workers, fishermen, the self-employed, unpaid family workers and apprentices. Furthermore, only workers between 14 and 65 years aged were included in sample. The final sample includes 168095 observations.

The estimation results are based upon six economies (3 years \times 2 regions). The two regions are Lisbon and the Tagus Valley on one hand and rest of the country on the other. Mean values in each economy of the variables used to construct the skill and complexity indexes are in Table 1. LTV stands apart when compared with the rest of the country. With respect to the variables used to calculate the skill index, the proportion of workers with lower secondary (9 years of education) or more is higher in this region. On the other hand, it has a lower proportion of workers with preparatory education (6 years) or less when compared with the rest of the country. Although the proportion of workers having college is fairly low in Portugal, there exist a significant asymmetry between the two distributions. In 1982 1.8% of the workers in LTV had college as compared with 0.6% in the other regions. The numbers for 1992 are equal to 2.9% and 0.8%, a much more pronounced difference than a decade before. Experience is also higher for workers employed in LTV.

The job characteristics also reveal pronounced differences between the two regions being examined. The proportion of scientific jobs and white-collar professions is higher in LTV. To a large extent this mirrors differences in the industrial composition of employment. Services dominate in this region. Manufacturing has the highest share of the employment in the rest of the country. It is worth mentioning, however, that the type of manufacturing differs between regions. LTV produces commodities using a less labour intensive technology. Finally, the average firm size is larger in LTV.

The estimation results for the supply side of the model (equation 8a) are in panel A of Table 2. The dependent variable is the logarithm of gross hourly wages. Following Murphy and Welch (1990), a third order polynomial for experience is included. Other covariates include nine dummy variables for highest level education completed and a dummy variable for gender. It also includes an intercept term for each of the six economies. These variables explain 79% of the variation of the dependent variable. Furthermore, all coefficients have the expected sign and ranking. Mean values and standard deviations of the resulting q_{it} index are also in the same table (panel C). This index increased over time in the LTV economy. With respect to the other regions, q_{it} also shows an increasing trend but this evolution is more apparent after 1986. But most important, the index is systematically higher in LTV than in the other regions.

Estimation results for equation (8b) are in panel B of Table 2. Of particular importance is the coefficient on q_{it} . The values point to a systematic difference in the return to the human capital index between LTV and the rest of the country (LTV being 10% higher) and an increase in the return to skill in LTV. The return to skill in the other regions is very similar in 1982 and 1986 but increased in the period subsequent to 1992.

The estimation results for the demand side of the model (i.e., equation 12) are in Table 3. The regression for q_{it} (see panel A) includes seven dummies for occupations, six dummies for the job level, seven dummies for the industry affiliation, and the logarithm of firm size. Three separate restricted versions of this specification that excluded occupations, job level and industry dummies were also estimated. Each of these restrictions were rejected. Mean values and standard deviations for the resulting j_{it} index are also in the same table. The mean is higher in LTV than in other regions. It increased in LTV while in the rest of the country it shows a v-shaped pattern.

Finally, the estimation results for equation (13) are presented in panel B of Table 3. The estimates for q_1 and q_2 have the expected positive signs, the values being equal to 0.3471 and 0.1441, respectively. Both values are statistically significant at the 1% level. This yields an estimated value for the scaling factor between $\ln c_{it}$ and j_{it} (d^*) equal to 6.94. The value for the complexity dispersion parameter (g) equals 2.41. The main implication of the model of a high return to human capital being associated with similar workers being assigned to more complex jobs is born out by the data. If the return to skill is 10% higher in economy 2 than in economy 1, the human capital index is 4.15% lower in economy 2 than in economy 1 for jobs of complexity c .

It should be noted that a simple model for the return to human capital would not have worked. The LTV region has the largest supply of highly skilled workers, see Table 1. So, just running the regression where the return to human capital is explained by the relative supplies of skill groups would yield the counterintuitive result that a large supply of highly skilled produces a high relative wage for this group. The assignment model provides a methodology for comparing the differences in distributions of supply and demand. In Lisbon, the high demand for highly skilled workers outweighs their large supply.

The value of g of 2.41 compares to values of about 3.5 estimated for the Netherlands⁵ and of about 3.8 estimated for the United States.⁶ Though these estimates refer to only three countries, each for a limited time period, one can conclude from the information available by now that the

⁵Calculated from Teulings (1995a). The q -index there is normalised by the return to university education, which can be calculated from Table 1 as the mean for 1982 and 1988: $(0.77+0.48)/2 = 0.63$. The expected value of q is 1.34 (Table 1, mean of personal characteristics, 3.30, minus intercept for 1982, 1.96). Then, from Table 4 and accounting for the normalisation of the q -index, we have: $d \sum_m b_m E[q]^m / d E[q] = 7.8$. Teulings (1995b) argues that equation (21) in Teulings (1995a) should be estimated with the j_{it} -index as the variable to explain, and not $\ln W_{it} - \sum_m b_m E[q]^m$. This is expected to reduce the estimate of b_m by a factor R^2 . Hence: $\gamma = 7.8 \times 0.45 \approx 3.5$.

⁶See Teulings (1995b). The q -index is normalised by the standard deviation of the residuals, i.e. 0.40. The expected value of the index is 0.8. From Table 2 and accounting for the normalisation of the q -index, we have: $d \sum_m b_m E[q]^m / d E[q] = 3.8$.

complexity dispersion varies in a limited range, between 2 and 4. Moreover, substitutability appears higher in Portugal than in the Netherlands and in the United States.

We did some further testing of the model. In particular we analysed whether the model covers intertemporal and regional variation in the same way. A simple test is to include a dummy for the LTV area in equation (13). Ideally, this dummy should be estimated to be zero, since $\ln a_{it}$ should cover all the variation in the allocation of worker- to job types between 6 economies (2 regions \times 3 years). If not, the coefficient q_2 , which covers only the intertemporal variation with the dummy for LTV included in the regression, should be positive in order to satisfy the second order condition. The model fails on both tests. The LTV dummy is clearly significant and the coefficient q_2 for the remaining intertemporal variation in our sample of 6 economies is small but negative (significantly). Hence, the model is capable of explaining the differences in urban versus rural returns on human capital, but cannot explain the rising trend in this return. One explanation might be that our assumption of technology within a job type to be constant is violated. When the job title remains equal but technology changes in a direction requiring better skilled workers, the data understate the increase in job complexity. We have no means for investigating this issue in further detail.

Applying the formula (6) for the log-cost increase for out-of-equilibrium wages setting for the case of a constant complexity dispersion parameter yields the following result:

D_w	log cost disadvantage
-1.00	3.205
-.050	0.470
-.020	0.056
-0.10	0.013
0.10	0.011
0.20	0.041
0.50	0.209
1.00	0.622

Hence paying a too low wage is more harmful than paying a too high wage. Better let a manager do the unskilled job than let the unskilled worker run the company. Paying half of the optimal wage is about to double the cost per efficiency unit of labour.

5 Conclusions

The modernization process in the Portuguese economy has been speeded up. The capital Lisbon and the surrounding Tagus Valley were the centre of this modernisation. The share of this region in the total population increased rapidly. The average education level of the work force in this region is above the nation wide average. This paper examines the forces driving this process. In particular, to what extent the supply of new more skilled generations sparks off the introduction of new technologies, or whether the process is demand driven with the new high tech economy pulling highly skilled workers to the LTV region.

The paper develops a model to analyse this process formally. We applied an assignment model of heterogeneous workers to heterogeneous jobs. The theoretical model establishes a link between the return to skill and allocation of workers to jobs. A simple estimation methodology of the model has been proposed. The main implication, a high return to skill being associated with similar workers being assigned to more complex jobs, is born out by the data.

The higher the return to the human capital index in an economy, the more precious will be highly skilled workers and the lower will be the skill level of the workers hired in a particular job type. The return on human capital increased over time but is always higher in the LTV region, as compared with the rest of the country. Despite the larger supply of highly educated workers in Lisbon relative to the other regions, their relative wages were higher because large supply was not enough to meet the even higher demand. Equally skilled workers get more complex jobs in

Lisbon compared to the countryside, basically because there are insufficient highly skilled workers to fill all complex jobs that are demanded in the high tech economy emerging in Lisbon area. Apparently, new technology in LTV region triggers the demand for human capital, not the other way around.

The estimation procedure allows us to obtain a simple, dimension free parameter, that measures the heterogeneity in jobs and therefore the ease of substitution between different worker types. The more heterogeneous are job types, the less easy is the substitution between worker types, see Teulings (1995b). The higher the variation in the assignment of worker types relative to the variation in return to skill, the smaller will be this parameter. Results based upon this parameter suggest that for a firm to pay half the optimal wage level and thereby attracting less skilled workers is about to double the cost per efficiency unit of labour.

Table 1 - Means of the variables used in the skill and complexity indexes by economy

	all	year82 Lisbon	year86 Lisbon	year92 Lisbon	year82 oth.regi.	year86 oth.regi.	year92 oth.regi.
<i>skill index</i>							
levels of education:							
less than primary	.0833	.1038	.0750	.0434	.1256	.0954	.0547
primary	.5508	.5170	.5045	.4195	.6533	.6510	.5632
preparatory	.1455	.1286	.1438	.1668	.1019	.1235	.2070
lower secondary	.0803	.0776	.0985	.1379	.0399	.0523	.0771
upper secondary (acad.)	.0529	.0494	.0632	.1065	.0200	.0277	.0523
upper secondary (commerce)	.0317	.0468	.0432	.0323	.0284	.0215	.0156
upper secondary (industry)	.0241	.0394	.0360	.0308	.0160	.0119	.0083
upper secondary (others)	.0057	.0043	.0048	.0159	.0017	.0031	.0051
intermediate courses	.0051	.0059	.0058	.0072	.0034	.0038	.0042
baccalaureate	.0068	.0089	.0086	.0106	.0044	.0038	.0045
college	.0139	.0183	.0167	.0290	.0056	.0059	.0079
experience	24.80	26.12	26.40	25.33	24.07	23.76	22.88
experience squared/10 ²	7.678	8.376	8.453	8.017	7.297	7.116	6.692
experience cubed/10 ⁴	2.740	3.067	3.061	2.878	2.594	2.495	2.290
female	.3148	.2547	.2796	.3246	.3038	.3498	.3853
<i>complexity index</i>							
scientific, tech. & artistic (c.0/1)	.0402	.0548	.0473	.0520	.0272	.0257	.0324
managerial & administrative (c.2)	.0173	.0224	.0215	.0344	.0068	.0083	.0104
clerical & related staff (c.3)	.2135	.2567	.2960	.2766	.1446	.1509	.1469
sales workers (c.4)	.0661	.0703	.0688	.0785	.0614	.0588	.0590
security & domestic services (c.5)	.0759	.0921	.0895	.1015	.0502	.0548	.0653
transport & production workers (c.7)	.1821	.0961	.1075	.1131	.2608	.2684	.2600
idem (codes 8)	.1654	.1551	.1435	.1265	.1907	.1865	.1904
idem (codes 9)	.2395	.2525	.2259	.2173	.2582	.2466	.2357
job level 1	.1122	.1126	.1054	.1043	.1143	.1224	.1143
job level 2	.2349	.1978	.2146	.1854	.2707	.2804	.2641
job level 3	.5004	.5005	.4921	.4779	.5166	.4982	.5161
job level 4	.0515	.0609	.0672	.0841	.0323	.0303	.0340
job level 5	.0548	.0681	.0617	.0670	.0437	.0425	.0441
job level 6	.0232	.0264	.0268	.0469	.0120	.0141	.0140
job level 7	.0230	.0336	.0320	.0345	.0104	.0122	.0134
mining & quarrying	.0075	.0051	.0043	.0044	.0102	.0110	.0105
manufacturing	.4793	.3677	.3465	.2911	.6248	.6484	.6123
electricity and gas	.0165	.0292	.0356	.0308	.0004	.0001	.0001
construction	.0889	.1042	.0762	.0909	.1008	.0768	.0845
wholesale, retail trade & hotels	.1900	.2239	.1989	.2291	.1642	.1543	.1670
transport, storage & communications	.0983	.1317	.1859	.1687	.0315	.0323	.0319
banking, insurance	.0694	.0922	.0988	.1159	.0340	.0336	.0394
community, social & personal services	.0500	.0460	.0538	.0691	.0341	.0435	.0543
log firm size	5.459	6.093	6.367	5.844	4.864	4.879	4.579

Note: c.=occupational codes; Lisbon=Lisbon and the Tagus Valley; oth. regi.=other regions

Table 2 - OLS estimation results for the supply side

A. dependent=log hourly earnings			B. dependent=log hourly earnings				
year82×Lisbon	3.7762	(.0083)*	year82×Lisbon	3.7896	(.0076)*		
year86×Lisbon	4.5485	(.0084)*	year86×Lisbon	4.4552	(.0081)*		
year92×Lisbon	5.3200	(.0085)*	year92×Lisbon	5.1712	(.0086)*		
year82×oth.regions	3.5605	(.0082)*	year82×oth.regions	3.6646	(.0082)*		
year86×oth.regions	4.2813	(.0083)*	year86×oth.regions	4.3903	(.0080)*		
year92×oth.regions	5.0640	(.0083)*	year92×oth.regions	5.0986	(.0077)*		
primary	.1647	(.0038)*	q _{it} ×(year82×Lisbon)	.9873	(.0070)*		
preparatory	.4164	(.0045)*	q _{it} ×(year86×Lisbon)	1.0867	(.0072)*		
lower secondary	.6911	(.0051)*	q _{it} ×(year92×Lisbon)	1.1327	(.0074)*		
upper secondary (acad.)	.8853	(.0058)*	q _{it} ×(year82×oth.regions)	.8860	(.0086)*		
upper secondary (commerce)	.8042	(.0064)*	q _{it} ×(year86×oth.regions)	.8808	(.0083)*		
upper secondary (industry)	.7087	(.0071)*	q _{it} ×(year92×oth.regions)	.9636	(.0077)*		
upper secondary (others)	.8626	(.0130)*					
intermediate courses	1.0036	(.0136)*	adjusted R ²	.7865			
baccalaureate	1.2593	(.0119)*	s.e.e	.3774			
college	1.4666	(.0089)*	F-statistic	56320			
			N	168095			
experience	.0590	(.0009)*					
experience squared/10 ²	-.1188	(.0034)*					
experience cubed/10 ⁴	.0713	(.0040)*					
female	-.2376	(.0020)*					
adjusted R ²	.7854						
s.e.e	.3785						
F-statistic	32390						
N	168095						
C. Means and standard deviations of the index q _{it}							
	all	year82 Lisbon	year86 Lisbon	year92 Lisbon	year82 oth.regi.	year86 oth.regi.	year92 oth.regi.
mean	1.0055	1.0518	1.0766	1.1218	.9130	.9148	.9489
standard deviation	.3061	.3086	.3020	.3225	.2690	.2761	.2885

Absolute standard errors in parentheses * denotes significant at the 1% level;

Table 3- OLS estimation results for the demand side

A. dependent= q_{it}			B. dependent= j_{it}				
year82×Lisbon	.9553	(.0075)*	intercept	-.0730	(.0010)*		
year86×Lisbon	.9732	(.0076)*					
year92×Lisbon	1.0083	(.0076)*	q_{it}	.3471	(.0010)*		
year82×oth.regions	.9243	(.0075)*					
year86×oth.regions	.9281	(.0075)*	$\log \alpha_{1t}$.1441	(.0032)*		
year92×oth.regions	.9591	(.0076)*					
managerial & administrative (c.2)	-.0565	(.0059)*	adjusted R ²	.4735			
clerical & related staff (c.3)	-.0196	(.0033)*	s.e.e	.1163			
sales workers (c.4)	-.0623	(.0039)*	F-statistic	75597			
security & domestic services (c.5)	-.1869	(.0038)*	N	168095			
transport & production workers (c.7)	-.2508	(.0034)*					
idem (codes 8)	-.1239	(.0034)*					
idem (codes 9)	-.1373	(.0033)*					
job level 2	.0165	(.0021)*					
job level 3	.1338	(.0020)*					
job level 4	.3218	(.0031)*					
job level 5	.3314	(.0031)*					
job level 6	.4689	(.0042)*					
job level 7	.7346	(.0056)*					
manufacturing	-.0646	(.0063)*					
electricity and gas	.0329	(.0077)*					
construction	-.0555	(.0065)*					
wholesale, retail trade & hotels	-.0410	(.0065)*					
transport, storage & communications	-.0308	(.0066)*					
banking, insurance	.1068	(.0067)*					
community, social & personal services	-.0534	(.0068)*					
log firm size	.0143	(.0003)*					
adjusted R ²		.4746					
s.e.e		.2219					
F-statistic		5840					
N		168095					
<i>hypotheses testing:</i>							
occupations irrelevant		F=2408					
job level irrelevant		F=6803					
industries irrelevant		F= 696					
C. Means and standard deviations of the index j_{it}							
	all	year82 Lisbon	year86 Lisbon	year92 Lisbon	year82 oth.regi.	year86 oth.regi.	year92 oth.regi.
mean	.2739	.3125	.3179	.3259	.2274	.2259	.2283
standard deviation	.1603	.1688	.1657	.1744	.1329	.1351	.1372

Absolute standard errors in parentheses * denotes significant at the 1% level

References:

- Kahn, S. and Lang, K. (1988) "Efficient Estimation of Structural Hedonic Systems", *International Economic Review*, **29**, 157-166.
- Murphy, K. and Welch, F. (1990) "Empirical Age-Earnings Profiles", *Journal of Labor Economics*, **8**, 202-229.
- Sattinger, M. (1975) "Comparative Advantage and the Distribution of Earnings and Abilities", *Econometrica*, **43**, 455-468.
- Teulings, C. N. (1995a) "The Wage Distribution in a Model of the Assignment of Skills to Jobs", *Journal of Political Economy*, **103**, 280-315.
- Teulings, C. N. (1995b) "A Generalized Assignment Model of Workers to Jobs for the US Economy, Discussion Paper, TI-95-67, Amsterdam: Tinbergen Institute.

figures to be inserted in the text

Fig 1. -

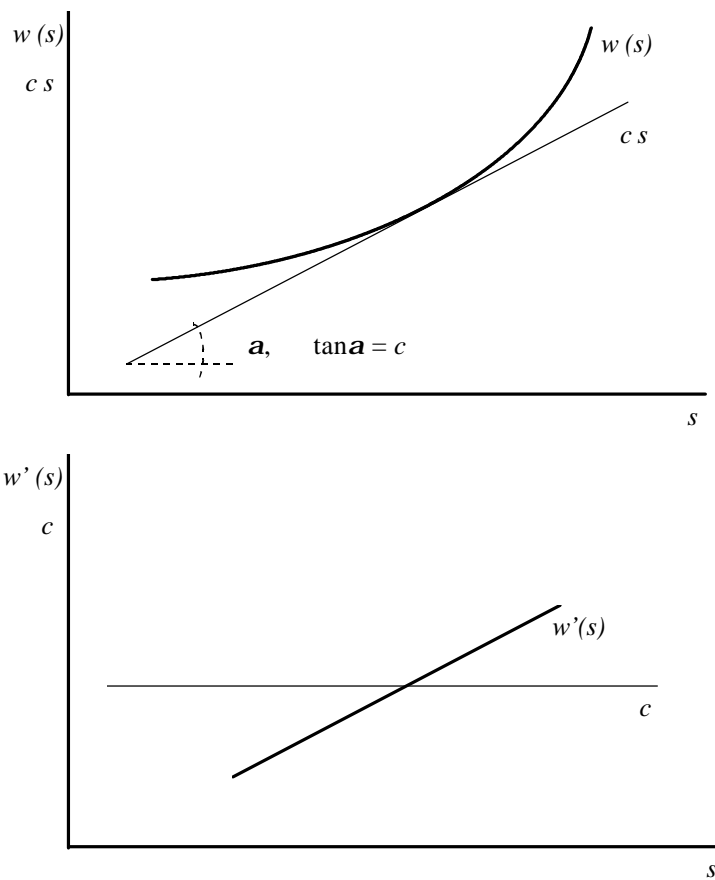


Fig. 2 -

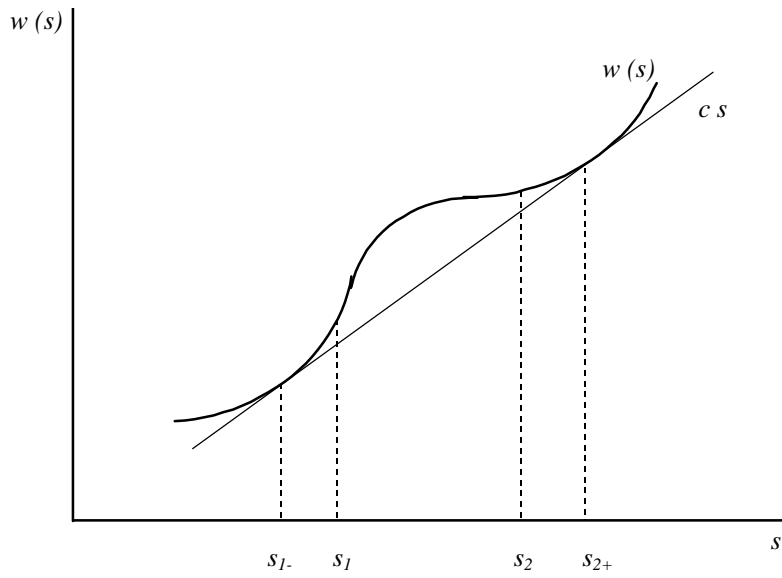


Fig. 3 -

