TITLE:
Best Techniques Leave Little Room for Substitution. A New Critique of the Production Function

AUTHORS:
Götz Kersting, Bertram Schefold

CORRESPONDING AUTHOR:
Bertram Schefold
Fachbereich Wirtschaftswissenschaften, RuW Fach 70
Johann Wolfgang Goethe-Universität
Theodor-W.-Adorno-Platz 4,
60323 Frankfurt am Main,

Tel home +4969 555 371
Tel office +4969 798 34759
Fax +4969 798 34544

schefold@wiwi.uni-frankfurt.de

Best Techniques Leave Little Room for Substitution. A New Critique of the Production Function

Götz Kersting
Institut für Mathematik
Johann Wolfgang Goethe-Universität*

Bertram Schefold
Fachbereich Wirtschaftswissenschaften, RuW Fach 70
Johann Wolfgang Goethe-Universität†

March 26, 2021

Abstract

Samuelson assumed a linear wage curve for each of a continuum of techniques such that their envelope was a monotonically falling wage curve for the economy, from which an aggregate production function fulfilling the marginal productivity conditions could be derived. But the capital intensities of the techniques chosen at each rate of profit are not necessarily lower at higher rates of profit, if the wage curves are not linear, a possibility exemplified by reswitching. This critique of the capital controversy does not rule out Samuelson's construction as an approximation, since the paradoxes have been shown to be rare. Instead, a possibility is likely that has so far not been noticed: the envelope of the wage curves will in the relevant range of the rate of profit be dominated by a small number of efficient techniques of approximately equal capital intensity, leaving little room for substitution. A new mathematical theorem demonstrates that the expected number of techniques that appear on the envelope is given by $(2/3)\ln s$. Numerical experiments and empirical investigations confirm the analysis.

Key words: Capital theory, production function, substitution, reswitching
B24, C62, C67, D57

1 Introduction

Paul A. Samuelson's article Parable and Realism in Capital Theory: The Surrogate Production Function, published 1962 in the Review of Economic Studies (Samuelson 1962), marked the beginning of what was probably the most conspicuous controversy in economic theory in the second half of the 20th century (Harcourt 1972; Kurz/Salvadori 1995). Only a few recall it today, although the discussion never ceased altogether (Kurz 2000; Hagemann 2020; Kurz 2020; Schefold 2020; Weizsäcker 2020). A brief summary of the debate seems indispensable before we come to what we regard as a very surprising new turnaround. It has been shown recently that the

*Robert-Mayer-Str. 10, 60325 Frankfurt am Main, kersting@math.uni-frankfurt.de
†Theodor-W.-Adorno-Platz 4, 60323 Frankfurt am Main, schefold@wiwi.uni-frankfurt.de

pathologies of neoclassical capital theory, on which the early critiques were based, are quite rare in economies with many sectors (Schefold 2017), compared with the two or three sector models (Petri 2011; Mainwaring and Steedman 2000), which were predominantly used in the first phase of the debate. Even readers unfamiliar with it may remember that it had revealed the possibility of reverse capital deepening (Levhari 1965; Samuelson 1966). This means a change of technique, induced by a change of the rate of interest, such that an increase of this rate is associated with a higher intensity of capital and, similarly, a lowering of the real wage may induce the switch to a technique, which is less, not more labour-intensive. These anomalies can now be shown to be rare in large economic systems, both theoretically (Schefold 2016; Schefold 2018) and empirically (Han and Schefold 2006). If they were frequent, they would not only in theory, but also in applied economics question the working of the substitution mechanism on which the determination of equilibrium through supply and demand in capital and labour markets is based according to neoclassical theory. The new critique, to be presented below, concerns not only theory, but also economic policy. It predicts that the substitution possibilities are surprisingly few, even if the number of potential techniques in the given technology or "book of blueprints" is very large. It turns out that the intensity of capital is likely to stay virtually constant over the entire relevant range of variation of either the wage rate or the rate of profit. "Substitution possibilities" here refers to the efficient techniques that show up on the envelope of all wage curves, on the factor price frontier. A multiplicity of techniques and hence of substitution possibilities below the envelope remains, but they are inefficient. The new finding may thus help to explain the fact that major technologies stay in place, when the rates of wages and of interest change during the cycle, while minor activities may be affected: it may be easier to get personal for cleaning or for repairing cars, when unemployment rises and wages are depressed, but major technical changes follow from progress, not from shifts in distribution. The result is not due to rigidities or deviations from rational behaviour but follows from profit maximisation under conditions of perfect competition.

We shall first summarize the debate on capital theory; the notes summarize the essential mathematical background (Section 2 and Notes 1 – 7). We shall then give an intuitive explanation of why there is little substitution and little change of the capital-labour ratio in the relevant range of the rate of profit (Section 3). Section 4 provides an exact mathematical statement of the main result, supported by experimental evidence. Section 5 discusses possible objections and extensions, Section 6 buttresses the main argument by numerical experiments and empirical investigations. Readers interested in the main mathematical result can go directly to Section 4, after having familiarized themselves with the problematic.

2 The capital theory debate

We start with a summary description of the origin of the debate. The marginal productivity theory of distribution had been controversial since its inception in the 1870s; Joan Robinson (1953–1954) and Piero Sraffa (1960) had criticized the concept of aggregate capital as its logical basis. Now Samuelson (1962) tried to defend it by deriving the aggregate production function from the set of possible steady states in a linear activities model.

To each level of distribution between labour and capital, characterized by the rate of interest, there corresponded a wage rate w, given the technique[1]. The factor price frontier $w(r)$ was monotonically falling and the same construct as the wage curve in Sraffa[2] (1960). We here refer to the Notes for formal details, which are well known to all connoisseurs of the critique so that references are not necessary. Samuelson succeeded in showing that a value of capital[3], equal to the value of the capital goods used with the given technique at the steady state prices, could be calculated at each level of r between zero and a maximum rate of profit or interest R. The efficient techniques[4] (competitive and corresponding to the maximum attainable level of w) would change with the rate of profit such that the capital-intensity would increase, as the rate of interest was lowered (corresponding to the fundamental neoclassical proposition that lowering the rate of interest facilitates investment and accumulation), and in such a way that lowering the wage rate would favour the adoption of labour-intensive techniques (corresponding to the fundamental neoclassical proposition that wages need to be lowered to increase employment). The envelope of the wage curve would express the same as a production function (see below Note 7).

Samuelson had assumed, as he himself had pointed out repeatedly, that the capital-intensities of the different sectors were the same in all industries of any given technique, so that the individual wage curves of each technique were straight lines[5]. But it was found in the subsequent controversy that the individual wage curves were not straight, their envelope, though falling, was therefore not necessarily convex; it could contain concave parts, hence the aggregate value of capital did not necessarily fall relative to labour employed, as the rate of interest or profit was increased, neither for the individual techniques, nor on the envelope. To the extent that this happened without a choice of technique and was due to the curvature of the individual wage curves on the envelope, these occurrences were called (non-neoclassical) Wicksell-effects. Even more striking was the phenomenon of reverse capital deepening. As the rate of interest increases, techniques change at the intersection of individual wage curves on the envelope. These are the so-called switch-points on the envelope, where generically one method changes so that there are two systems of prices, differing in one method; prices are the same for both systems at the switch-point itself. It turned out to be possible that a technique, adopted because of a rise of the rate of interest across the switch-point, showed a higher, not a lower intensity of capital and indeed, a technique, which had been profitable at a low rate of profit and which was dominated by other methods at intermediate profit rates, could reappear at a high rate (so called reswitching)[6]. Samuelson's student Levhari (1965) tried to exclude reswitching by means of assuming that each technique was given by an indecomposable input-output matrix, but his proof was erroneous, counter-examples were provided by a number of critics, starting with Pasinetti (1966), and the story of this battle was told in quite a few histories, short and long (Harcourt 1962; Kurz and Salvadori 1965; Hagemann 2020). After this, advanced theorists like Samuelson himself admitted that the use of the aggregate production function was not rigorous (Samuelson 1966), but it did not disappear from the textbooks and returned with the new theories of growth in the 1980s (Aghion and Howitt 1998).

Garegnani (1970), in what was perhaps the most thorough contribution to the critique, tried to show that Samuelson's assumption had, by being restrictive, not only been sufficient to show that the surrogate production function could be constructed, but Garegnani tried to demonstrate that these assumptions were also necessary, and not only for the validity of neoclassical theory in the form of the production function (the Clarkian parable) but also for Walras's general equilibrium (as a theory of long-run equilibrium) and for Böhm-Bawerk's analysis. As we discuss in the Notes[7] and partly in other papers (Schefold 2013a), Garegnani's proof of the necessity omitted a small possibility in the case of the aggregate production function and was more tentative than rigorous as regards the 19th century neoclassical economists. What was on that occasion in 1970 more important, he failed to clarify the relevance of his critique for modern general intertemporal equilibrium, which is normally not a long-run equilibrium because the initial endowments are not in general given in that proportion which would allow to reach a steady state from the beginning and could then be maintained for subsequent periods. The question as to how the critique related to intertemporal equilibrium had been raised by the very editor of the Journal

(Bliss 1970). Garegnani endeavoured to answer to this challenge in later years, and this entailed a controversy of its own, which we cannot pursue here (Schefold 2008). Suffice it to say that non-neoclassical Wicksell-effects and reverse capital deepening imply a special kind of stability problems for general intertemporal equilibria, even if their existence is not in question.

By contrast, Garegnani's general assertion, that the critique affected all versions of the old long-run neoclassical theory, was rooted in his early work (Garegnani 1960) and turned out to be a profound insight. It has been developed later by Petri (2004), Schefold (2016), among others, and Garegnani's chief merit was perhaps his identification of the break in the history of neoclassical thought, when early long-run neoclassical theory ended, because Lindahl and Hicks opted successfully for the introduction of temporary equilibria (Garegnani 1976), with the result that the long-period method was largely abandoned by neoclassical theorists in favour of the intertemporal approach, except for the steady states, which were kept as a framework to discuss theories of growth (e. g. von Weizsäcker 1971).

But, in all this, nobody seemed to question or even to notice the least plausible of Samuelson's assumptions. What was noticed, and pointed out by Samuelson himself, was that the individual wage curves had to be straight lines, for then the capital-intensity was constant for each technique and hence an unambiguous characteristic of the wage curve: each technique had its specific capital-intensity. For the wage curves to be straight, the capital-intensities had to be the same in all sectors, as already stated, and this means that relative prices did not depend on distribution and therefore were equal to labour values. As Salvadori and Steedman (1988) have shown, methods of production taken from different techniques could thus not be combined, for the combinations would have led to unequal capital intensities in the different sectors, and the straight wage curves would at the intersections on the envelope have been dominated by the wage curves of combinations, which would thus in general not be straight[8]. Samuelson himself had very briefly indicated that the capital goods in his model were specific to the techniques. This means that each technique produces and uses capital goods that are not produced or used in other techniques, while we shall assume that methods taken from different techniques can be combined to form a new technique (but we shall discuss another restriction on technological choice as an intermediate case in Section 5). The critics have quite correctly pointed out the lack of generality of Samuelson's construction[9], which is expressed in these properties, and they have noted the irony that Samuelson based the surrogate production function on the assumption that the labour theory of value held for each of the alternative techniques, as in the first two volumes of Das Kapital by Karl Marx, while he wanted to distance himself from the Marxian tradition (Samuelson 1966 [1961]).

But all the critics failed to perceive that Samuelson did not only assume that individual wage curves were straight, but also that wage curves of techniques with a high productivity of labour predominantly had a low maximum rate of profit or, with constant returns to scale, a low maximum rate of growth, equal to the maximum rate of profit. In his diagram, the techniques can be seen to be ordered according to the productivity of labour, equal to $w(0)$, the real wage per head, if all output goes to workers. Each technique is also characterized by a maximum rate of profit, and these two points determine the straight wage curve. Samuelson's assumption, as shown in his diagram, but not formulated in words, is that, as in Diagram 1, Section 3 below, the permutation of the maximum rates of profit associated with the technique is the inverse ordering of the ordering of the techniques according to the productivity of labour – except perhaps for a few inferior techniques, which do not get up on the envelope. If it is not assumed that the ordering of the maximum rates of profit is inverse to that of the productivities of labour or if not some similar assumption is made, not a few but most techniques will have wage curves below the envelope and the number of the techniques on the envelope will be so small that the idea of substitution (which is what all the discussion is about) loses its meaning. This is what we shall show.

3 From reswitching to the number of substitution possibilities

How many possibilities of substitution do we encounter on the envelope? In order to clarify the question, it is useful to ask why a technique should not be 'good' both with regard to the productivity of labour and the maximum rate of growth. This question presumably never came up in the discussion, because neoclassical economists repressed the entire discussion in their minds and regarded the idea of substitution as obvious, while the critics thought that the wage curves of individual techniques were curves anyway, which formed a complicated pattern such that most of them came up on the envelope, on the factor price frontier, in unpredictable sequences. I (here Bertram Schefold is writing) can claim an exception for myself. It is now (spring 2020) exactly 50 years ago that I asked Joan Robinson in her class in Cambridge, whether she expected 'many' wage curves on the envelope. Sraffa himself had spoken of a "rapid succession of switches" (Sraffa 1960, p. 85), which one would encounter as one moved down the envelope. She replied that she thought there would be only one wage curve dominating the others, that of the 'best' technique; it would be superior independently of the level of distribution. I was astonished. We often discussed the choice of techniques and switch-points, when we experimented with two-sector models (the calculations of prices derived from input-output models came later). When I asked her about visible changes of technique, she replied that that was technical progress. I now had a doubt in my mind, but I returned to my agenda, which was the theory of joint production.

In the early 2000s, a Korean student of mine, Zonghie Han, wanted to find empirical proofs for the existence of reswitching and reverse capital deepening. He examined the envelopes of wage curves, calculating the individual wage curves from empirical input-output tables, which, taking in combination as books of blueprints, gave rise to a rich choice of techniques. In order to keep the calculations manageable, he would always combine the input-output tables of two countries at one time or of one country at two different times, assuming that in each industry there was the choice between two methods, say, either car manufacturing French style or German style. Since the input-output tables had 33 relevant sectors, there were now 2^{33} potential techniques for the two countries taken together. If all the methods of production had been assumed to be specific for the technique in which they were used, and that means, in this example, for the country concerned, only two wage curves, one of France and one of Germany, would have been to be compared. Intermediate cases of specificity are plausible: Both Austria and Greece produce wine, but only Greece has maritime transportation. Since Han ruled out such specificity completely (which was of course as daring an assumption as the opposite)[10], two input-output tables gave rise to the full potential of a spectrum of 2^{33} different wage curves. The envelope of each spectrum was calculated by linear programming, and pairing a number of input-output tables in this manner, several hundred envelopes were obtained, which Han inspected painstakingly. A thesis resulted (Han 2003) and a joint paper (Han and Schefold 2006) containing two surprises: (1) Cases of reverse capital deepening and one case of reswitching could be found, but they were rare. More than 95% of the switches were neoclassical. (2) About ten wage curves of individual techniques formed each envelope; not one, as Joan Robinson had anticipated, and not many, as Samuelson and Sraffa had implied, without making their assumptions explicit – indeed, they did not even realize that they here were assuming something.

I (BS) have spent much time during the last fifteen years trying to confirm, reject or at least to understand these results. Before we get to puzzle (2), which is the subject of this paper, pointedly reformulated, we must summarise some results regarding (1).

Numerous studies show that wage curves derived from input-output tables are never straight, but they are not as strongly curved as the critics used to expect. The reasons are not entirely clear. It is well known that a linear wage curve results if prices are expressed in terms of baskets of goods as numéraires and if the basket is proportional to Sraffa's standard commodity[11]. Appropriate measures for prices in such empirical comparisons are not individual goods – they, taken as numéraires, lead to more pronounced curvatures in most cases – but an average of the vectors of consumption or of the vectors of net national products in the comparison of the wage curves derived from the input-output systems of different countries in a given period. These vectors will in general not be equal to the standard commodity of any of the countries concerned, but their compositions may be somewhat similar, so that the deviation of the wage curves from linearity may be moderate[12].

It is also known that linear wage curves result, if prices are equal to labour values (this, strictly assumed, was Samuelson's case). Actual prices deviate from labour values, but perhaps not that much, as already Ricardo thought[13]. Hence another reason why the deviation of actual wage curves from linearity may be moderate, and the two reasons given do not exclude, but may reinforce each other.

A third possibility was investigated on purpose in order to explain the phenomenon of near linearity of wage curves in the context of the critique of the surrogate production function. If the systems are random in that the input-output matrices are random (with certain additional properties) and if the labour vector stands in a certain random relationship to the matrix, a nearly linear wage curve will result, and it tends to strict linearity as the dimension of the system (the number of sectors) tends to infinity. These assumptions have been used to construct an approximate surrogate production function[14], and they may also explain the existence and the stability in the factor markets of general equilibria of the "old" 19th-century neoclassical authors[15].

So there are reasons to return to Samuelson's reconstruction of neoclassical theory on the basis of linear wage curves – not in the sense of a rigorous theory that holds without exception, but as an approximation, this time not based on the labour theory of value, but on weaker and perhaps more plausible assumptions. Admittedly, linear wage curves are a fiction. Theory and empirical investigations leave no doubt that Wicksell-effects are ubiquitous, even if the deviation from linearity is, especially for numéraires consisting of many commodities, not large. The examination of this fiction remains nonetheless essential, for it is necessary for neoclassical theory to avoid declination and reverse capital deepening. We do, of course, not pretend that wage curves are linear, but we propose to return, temporarily at least, to the mental experiment of linear wage curves. This should be of interest also for readers who are not inclined to regard empirical wage curves as quasi-linear. As will be seen, the assumption is not essential for all of what follows. We shall argue in Section 5 that weaker assumptions suffice to preserve essential results.

Cutting a long story short, we now take up point (2) and assume strictly linear wage curves: what then about Samuelson's assumption that, the higher the labour productivity of an individual technique, the lower the corresponding maximum rate of profit?[16] Samuelson seems to have taken it for granted that, by and large, for any given technique, the productivity of capital would be the lower, the higher the productivity of labour. If this is the case, we shall say that we have a spectrum of techniques with inverse productivities. That spectra are not always like that, follows from the following consideration: Joan Robinson thought that "good" techniques would be characterized by high productivities of both, of labour and of capital. If there is a really good technique, it dominates all others and there is no substitution. Samuelson's assumption implies that, as one moves down the envelope, all or most techniques will appear one by one. He gave a graphic example of techniques alpha, beta, gamma, delta, appearing in this order on the envelope, except that a process epsilon remained entirely below.

The number of wage curves involved here can be very large. In our standard example, we have ten countries, each with 100 industries or sectors, producing the same commodities. Information is decentralized, in that producers know or can find out what the technique of each of their nine competitors in their industry is, while nobody knows the peculiarities of all industries in all countries. If we assume away any possible specificity of the capital goods or of combinations of processes, ten methods can in principle be used in each industry, so that there are in principle 10^{100} wage curves, and the question now is: how many of them will make it up on the envelope? Most of them, as Samuelson's construction seems to suggest – then Sraffa's formulation of a "rapid succession of switches" is appropriate – or only one, because there is only one dominant technique independently of distribution, as Joan Robinson thought? Both positions have intuitive appeal. The neoclassical might say: if the productivity of labour is high, the engineers have probably achieved this at the expense of a high productivity of capital. Joan Robinson might reply that research is always directed at finding methods that beat those of the competitors in a broad set of circumstances, that entrepreneurs aim at surplus profits without favouring the use of capital or labour, that sunk costs of finding new techniques are high and that therefore successful techniques are not numerous and do not easily change because of redistribution.

The central assumption to begin with is in a sense a compromise: The two productivities shall be random and completely independent of each other. More formally: we can order each finite set of techniques $\sigma = 1\ldots, s$, with wage curves $w^\sigma(r)$, such that the ordering corresponds to the productivities of labour and $w^1(0) > \cdots > w^s(0)$. The assumption then is that the ordering of the corresponding $R_1, \ldots, R_s$ of the maximum rates of profits or productivities of capital shall be independent of the ordering of the productivities of labour or that, which is an equivalent expression of the central assumption, all permutations of the R_σ are equally likely. We thus write the ascending sequence of the maximum rates of profit as $R_{\sigma_1}, \ldots, R_{\sigma_s}$, where $(\sigma_1, \ldots, \sigma_s)$ is a permutation of $(1, \ldots, s)$. R_s is the maximum rate of profit pertaining to wage curve $w^s(r)$. R_{σ_s} is the largest among all s maximum rates of profit.

The assumption can be derived from the postulate of a uniform bivariate probability distribution, but it is more general, as we shall see. Meanwhile it turns out that the assumption leads to a new mathematical problem: how many of these s wage curves can be expected to appear on the envelope? The question seems not to have been asked before, except for a paper by Schefold (2013b), which gives an easy provisional answer. The complete solution, to be presented below, is more difficult. It has been found by Götz Kersting and has surprising implications.

We begin by summarizing Schefold's sketch of the matter. As we shall see, it leads to an upper bound, if we proceed "from above" and look at the wage curve with the highest productivity of labour, $w^1(r)$. It is on the envelope by definition, with $w^1(0) > w^\sigma(0)$; $\sigma = 2, \ldots, s$. Coming down from above, it is clear that $w^2(r)$ will be on the envelope with probability $\frac{1}{2}$, for it will be on the envelope if $R_2 > R_1$ and it will be dominated by the first wage curve throughout if $R_1 > R_2$. Both possibilities are equally likely by assumption (we assume all maximum rates of profit and all productivities of labour to be different). Continuing to look from above, we find that $w^3(r)$ appears on the envelope with probability $\frac{1}{3}$, for $R_3 > R_1$ and $R_3 > R_2$ is possible both with $R_1 > R_2$ and $R_2 > R_1$, hence in two cases out of six, hence with probability $\frac{1}{3}$. By induction, we find that the expected number of wage curves on the envelope, ω, is given by $1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{s}$. If s is large, we can approximate ω by the natural logarithm:

$$\omega = \ln s.$$

The natural logarithm of s tends to infinity, but slowly; the share of the wage curves, which appear on the envelope, Ω, is given by $\Omega = \frac{\omega}{s} = \frac{\ln s}{s}$, and this tends to zero as s tends to infinity; the number of efficient techniques becomes vanishingly small relatively to the number of potential techniques below the envelope.

It has been easy to derive this estimate, but it overstates the number of wage curves on the envelope, for the reason that a wage curve coming later, as one moves down on the envelope, can become dominant (appear on the envelope) by dominating one or even several earlier wage curves, which seemed to have become dominant by looking only from above. The possibility is illustrated in Diagram 1.

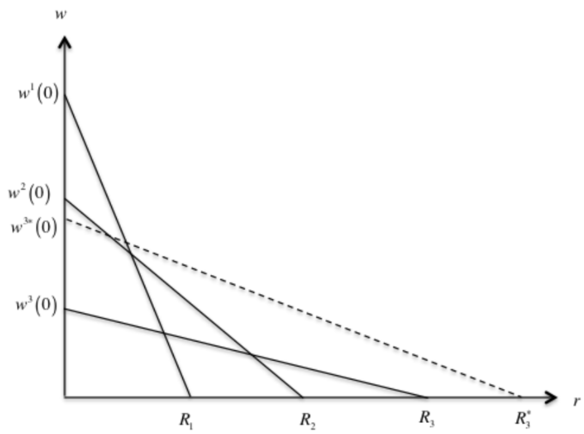


Diagram 1: Three wage curves with neoclassical ordering (the ordering of the maximum rates of profit is inverse to that of the wage rates). If the maximum wage of the third wage curve is higher (dotted line), one switch-point disappears (dominance from below).

The unbroken lines are three straight wage curves of the neoclassical type: the ordering of the productivities of capital is inverse to the ordering of the productivities of labour. But the dotted line shows a possible alternative to the third wage curve such that it dominates the second wage curve, which seemed to be on the envelope, when one looked from above and considered only the first two wage curves.

Because of the phenomenon of 'dominance from below', the formula $\omega = \ln s$ overstates the number of wage curves to be expected on the envelope, but by how much? This is a curiously tricky mathematical problem. It is clear that both $w^1(r)$ and $w^{\sigma_s}(r)$ can never be dominated from below, but a $w^{\sigma}(r)$, where σ is "in the middle", can possibly dominate many wage curves with higher labour productivities from below at intermediate rates of profit, which might have been dominant, looked at from above. Yet, the formula $\omega = \ln s$ seemed to provide an only moderate overstatement according to empirical investigations, applying the formula to the results in Han and Schefold (2006), in Zambelli (2017) and also in numerical experiments (some are reproduced below, Diagram 5). The numerical experiments were based on the assumption that the wage rates at $r = 0$ and the maximum rates of profit were at equal distances (the wage curves were drawn in a regular grid). Looking at Diagram 1, one realises that the transition to a regular grid can make a difference: if $w^{3*}(0)$ is shifted downwards, so as to obtain as much space between $w^2(0)$ and $w^{3*}(0)$ as between $w^1(0)$ and $w^2(0)$, the dominance from below disappears. But dominance from below can evidently also occur, if the grid is regular.

Proposition 1
Under the central assumption (equal probability of all orderings of the maximum rates of profit) and if s straight wage curves are given, the expected number of wage curves on the envelope is equal to or lower than $\ln s$.

The empirical confirmations of Proposition 1 can only be approximate: real wage curves, as derived from empirical input-output tables, are not exactly straight; numerical experiments with straight wage curves are based on regular grids. Nonetheless, the empirical results are plausible and the results will be extended in Sections 5 and 6.

The logarithm $\ln s$ tends to infinity with s, but slowly. In the standard example, with 10^{100} wage curves, the number of curves on the envelope is only $100 \ln 10 \cong 230$. The substitution possibilities are therefore far more limited than the enormous number of wage curves involved suggests. But we have found a much more radical result: there is virtually no change of the capital-labour ratio along the relevant range of the envelope, as w becomes large, and the possibilities of substitution cluster in two small neighbourhoods: around the maximum wage rate $w^1(0)$ and the largest of the maximum rates of profit R_{σ_s}. The exact formulation of these propositions requires some preparation and more specific assumptions, summarized in the Remark below.

Let the wage curves $w^1(r), \ldots, w^s(r)$, with $w^1(0) > \cdots > w^s(0)$ (straight lines) of a spectrum of techniques be arranged in the rectangle Q, the wage curve box. The line segments $w^\sigma(r)$, defined by the end points $w^\sigma(0)$ and R_σ, can be represented by a point P_σ in Q, with the end points as coordinates. Q contains a grid, the horizontal lines of which are defined by $w^\sigma(0)$ and the vertical lines by the R_σ. We measure profit rates on the abscissa, which are pure numbers (pure numbers apart from the time dimension of interest, here taken into account via the length of the period of production). We measure on the ordinate output per head. When we look at the grid, these measures define the distances between the lines of the grid. The lines themselves may simply be numbered. Since the $w^\sigma(0)$ and the R_σ are different, each horizontal and each vertical line will contain exactly one representative or anchor point P_σ, so that the grid in Q looks like a perturbation matrix, except that Q is not necessarily square and the grid is in general not regular, in that the distances between the grid lines vary. See Diagram 2, where we have five wage curves with the corresponding anchor points.

One easily confirms: Two (straight) wage curves like w^1 and w^2 in the wage curve box intersect inside the wage curve box if and only if the line segment $\overline{P_1 P_2}$ is negatively sloped. Dominance occurs if and only if $\overline{P_1 P_2}$ is positively sloped. Finally, w^1 and w^2 intersect at their endpoints on the boundary of Q, if and only if $\overline{P_1 P_2}$ is horizontal or vertical.

Now we imagine that the wage curve box gets below the diagonal d filled with more and more wage curves of number s, in such a way that all orderings of the maximum rates of profits are equally likely. We therefore stick to the central assumption adopted for Proposition 1. We can express it by saying that the pair $[w_\sigma(0), R_\sigma]$ is independent from the pair $[w_\tau(0), R_\tau]$, $\sigma \neq \tau$; for all σ and τ. The five wage curves drawn correspond to Samuelson's idea of a surrogate production function, but this picture will be destroyed if there is a wage curve dominating most of the envelope of w^1, w^2, w^3, w^4, such as a wage curve w^6 (dashed line, roughly parallel to diagonal d) and represented by point P_6. If one or several such points exist in the upper right corner of the box $ABFE$, say in the triangle DFG, a new envelope arises with – as will be shown – only a few switch-points except near the corners E and B. As we have drawn the diagram, the envelope consists of short stretches of w^1 and w^4, and of the long stretch of w^6. Substitutions, which will take place on this envelope, will only occur near the corners E and B, if there is only one such point P_6 in the triangle DFG. If there are several such points, the slopes of the

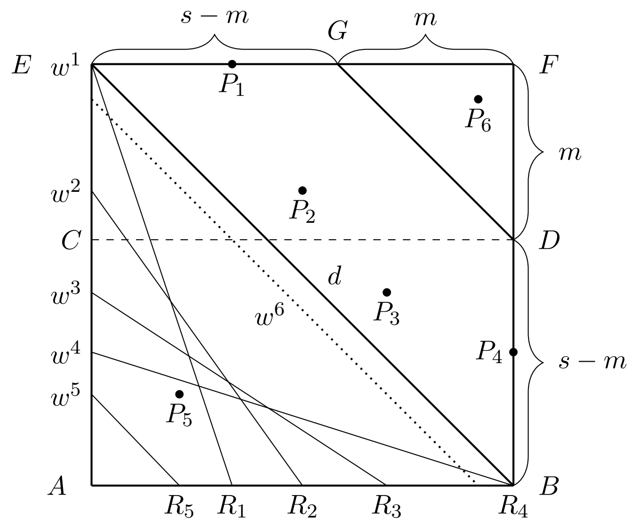


Diagram 2: The wage curves and their anchor by points in the wage curve box. One wage curve, w^5, is entirely dominated by the others, and such domination shows in the fact that P_1, P_2, P_3, P_4 are all above and to the right of P_5.

corresponding wage curves will necessarily be almost equal, hence substitutions would (except at the corners) not lead to appreciable variations of the capital-labour ratio.

How likely is that to happen? In order to show it, we calculate the probability π that it does not happen, that is, we calculate the probability π that there is no anchor point in the triangle DFG; therefore that all s anchor points lie in $ABDGE$. Triangle DFG is on m rows and m columns, trapezium $CDGE$ is on m lines with rows of lengths between $s-m$ and s. The point on the first row can therefore lie on one of $s-m$ places, the point on the second row (which, in trapezium $CDGE$, has $s-m+1$ places) can also lie on one of $s-m$ places, since one column has been occupied by the point on the first row. Similarly for $m-2$ later rows. The number of possible placements of points in the trapezium is $(s-m)^m$.

Since m columns have been occupied in the trapezium, there remain $s-m$ columns to be occupied in the rectangle $ABDC$ with $s-m$ rows; $s-m$ points can be placed in the remaining rectangle $ABDC$ in $(s-m)!$ ways.

The entire box $ABFE$ can be occupied in $s!$ ways. Hence the probability π we are looking for is, with $0 \le m < s$, given by

$$\pi = \frac{(s-m)^m (s-m)!}{s!}.$$

Two special cases may be considered, before we transform this formula:

If $m = 0$, trivially $\pi = 1$.

If $m = 1$, $\pi = \frac{s-1}{s} = 1 - \frac{1}{s}$ and $\bar{\pi}$, the probability that at least one point like P_1 is in the upper right corner of Q, is $\bar{\pi} = \frac{1}{s}$. This is the Joan Robinson case: One wage curve, the diagonal d, representing one best technique, dominates all others. This is unlikely, but the surprising result will be that the unlikely case becomes likely with only small modifications.

In order to calculate the general case, it is convenient to begin with $1/\pi$:

$$\begin{aligned}\frac{1}{\pi} &= \frac{s!}{(s-m)^m(s-m)!} = \frac{s(s-1)\cdots(s-m+1)}{(s-m)^m} \\ &= \frac{s}{s-m}\cdot\frac{s-1}{s-m}\cdots\frac{s-m+1}{s-m} \\ &= \Big(1+\frac{m}{s-m}\Big)\Big(1+\frac{m-1}{s-m}\Big)\cdots\Big(1+\frac{1}{s-m}\Big)\end{aligned}$$

The ratios $m/(s-m), \dots, 1/(s-m)$ are small for large s, given m. Hence they may be approximated by using the formula $\ln(1+x) \cong x$, and the approximation will be exact, if we let s and m go to infinity, but in such a way that the ratio $m/(s-m)$ tends to zero. This is achieved by assuming[17] $m = \sqrt{\gamma s}$, where $1 \leq \gamma \leq s$, since we had assumed $m \leq s$. The coefficient γ then is kept constant, as $s \to \infty$:

$$\begin{aligned}\frac{1}{\pi} &= \exp\Big[\ln\Big(1+\frac{m}{s-m}\Big)+\cdots+\ln\Big(1+\frac{1}{s-m}\Big)\Big] \\ &\cong \exp\Big[\frac{1}{s-m}(1+\cdots+m)\Big] = \exp\Big[\frac{m(m+1)}{2(s-m)}\Big] \\ &\cong \exp\frac{\gamma s+\sqrt{\gamma s}}{2(s-\sqrt{\gamma s})} = \exp\frac{\gamma+\sqrt{\gamma/s}}{2-2\sqrt{\gamma/s}} \xrightarrow{s\to\infty} \sqrt{e^\gamma}\end{aligned}$$

We get

$$\lim_{s\to\infty}\pi = \frac{1}{\sqrt{e^\gamma}}$$

and $\bar{\pi}$ tends to

$$\lim_{s\to\infty}\bar{\pi} = 1-\frac{1}{\sqrt{e^\gamma}}.$$

If $\gamma = 1$, we thus find that the wage curves will be dominated by at least one wage curve very close to the diagonal d in roughly one half of the cases[18]. And if we set a higher value for the constant γ, this will become virtually certain, as $\bar{\pi}$ will tend to 1, without, however, reaching that limit s, since γ must be kept bounded.

Remark
The probability distribution underlying this result can be formalized in different ways; we have begun with an elementary and intuitive approach[19]. We now provide a formal description of the strong case: We assume a bivariate conditional probability distribution of the anchor points $[w^\sigma(0), R_\sigma]$, which is uniform. The distributions of the $w^\sigma(0)$ and the R_σ on the axes are uniform and the pairs $[w^\sigma(0), R_\sigma]$ are independent. Moreover, the following conditions hold:

(i) The number of grid lines s, the values of $w^1(0)$ and R_{σ_s} and the distances between the grid lines are given.

(ii) No two methods of production are equal, hence no two values of the $w^\sigma(0)$ or of the R_σ shall be equal; there is one and only one anchor point on each grid line or row.

(iii) If a σ with $w^\sigma(0)$ is given and m grid columns are occupied, R_σ is on any of the remaining grid columns with probability $1/(s-m)$.

(iv) Similarly, if the roles of rows and columns are reversed.

The definition implies that an anchor point $[w^\sigma(0), R_\sigma]$ is in its position on a given line σ with probability $1/s$ and in any position with probability $1/s^2$, as there are s^2 positions in the wage curve box, as long as the position of the other anchor points are unknown. Each increase of s implies the introduction of a new grid, and it must be specified whether and how the maximum dimensions $w^\sigma(0)$ and R_σ and the distances between the grid lines change (see Corollary 3 below). It is striking that no such specification was necessary for Proposition 1; the assumption about the permutations sufficed.

Note that the formula for π, the probability that there is no anchor point in triangle DFG, can be calculated, using (iii), by observing that $[w^1(0), R_1]$ is in its position in the trapezium $CDGE$ with probability $\frac{s-m}{s}$, if all other positions are unknown, $[w^2(0), R_2]$, given R_1, is in its position with probability $\frac{s-m}{s-1}$ and so on; $[w^m(0), R_m]$ with probability $\frac{s-m}{s-m+1}$, and all these conditions are fulfilled simultaneously if

$$\pi = \frac{s-m}{s} \cdot \frac{s-m}{s-1} \cdot ... \cdot \frac{s-m}{s-m+1} = \frac{(s-m)^m (s-m)!}{s!}.$$

This is the same formula as the one obtained above.

Note also that the assumption of a uniform distribution of the anchor points implies what we had called the "central assumption", as can be shown by induction. As with the throw of two dice, the positions of R_1 and R_2 are independent, hence $R_1 > R_2$ and $R_2 > R_1$ are equally probable, $R_1 = R_2$ being excluded. Similarly, $R_{\sigma+1} > R_\tau$ and $R_{\sigma+1} < R_\tau$ are equally probable; $\tau = 1, ..., \sigma$. Since all permutations can be generated by sequences of permutations of two elements, all permutations are equally probable. The converse, however, is not true. If x_{ij}; $i, j = 1, ..., s$; is the probability for an anchor point to be found on row i, column j, if $x_{ij} = 1/i$ for $j = 1, ..., i$ and if $x_{ij} = 0$ for $j = i+1, ..., s$, we have $P(R_\sigma > R_{\sigma-1}, ..., R_\sigma > R_1) = \frac{1}{s}$, although the probability distribution is not uniform. We shall show in Section 5 and 6 how the assumptions of the strong case can be relaxed and modified without affecting the main results substantially. To begin with a uniform distribution of the anchor points between given bounds is appropriate, because it is the maximum entropy probability distribution.

Now we conclude from what we have proved about π and $\bar{\pi}$:

Proposition 2
If the number s of techniques can be increased indefinitely, to each probability π^*, $0 < \pi^* < 1$, a number γ can be assigned such that with probability π^* there will be a wage curve w^* with a distance to the diagonal d, measured horizontally and vertically, not larger than $m = \sqrt{\gamma s}$.

Corollary 1
The envelope of all wage curves will lie in between w^* and d.

Note that wage curve w^6 in Diagram 2 illustrates such a w^*.

Corollary 2
The relative distance, bounded from above by $m/s = \sqrt{\gamma/s}$, tends to zero.

Will this distance also tend to zero absolutely (and not only relatively), if we measure it as percentage points of rates of profit along the abscissa and in terms of output per head along the ordinate? This will evidently depend on how we model the increase in the number of techniques. It need not be the case, if additional techniques have higher output per head and higher rates of profit – the wage curve box then will grow larger –, but if additional techniques spring up in between those which are already there – if the growth of the box is bounded –, our grid becomes denser, not larger and the distance between neighbouring points on the axes shrinks at both ends. Obviously:

Corollary 3
If additional techniques come in without an increase of maximum output per head and of the maximum of the rates of profit, such that the R_σ cluster uniformly all along the abscissa and the $w^\sigma(0)$ along the ordinate, the distance between w^* and d tends to zero.

In order to observe the process of clustering for large, but finite s, we go back to Diagram 2. As s increases, additional techniques are brought in and the productivities of labour and capital increase, but not indefinitely in a finite world. So let us assume a given large s – we may again think of our standard example with 10^{100} wage curves. We measure the wage rate and the rate of profit in relative terms by normalizing $w^1(0) = 1$ and $R_{\sigma_s} = 1$. We may impose a degree of certainty c that there is at least one wage curve close to diagonal d (corresponding to w^6 in Diagram 2) by requiring that $\pi \leq c$, so that $\sqrt{e^\gamma} \geq 1/c$ or $\gamma \geq \ln(1/c^2)$. If we demand that such a wage curve exists with the expectation of a trillion to one, we have $c = 10^{-12}$ and $\gamma = \ln(1/c^2) \cong 55$.

This defines $m = \sqrt{\gamma s}$; the wage rate $w^6(0)$ of this wage curve w^6 therefore is in an interval I_w between a $\tilde{w} = 1 - m/s$ and $w^1(0) = 1$. The maximum rate of profit R_6 is in an interval I_R between a $\tilde{R} = R_{\sigma_s} - m/s$ and $R_{\sigma_s} = 1$. In the standard example, we have $m \cong \sqrt{55} \cdot 10^{50}$. The length of the intervals I_w and I_R[20] is $m/s \cong 7 \cdot 10^{-50}$. If s is larger than astronomical numbers in the standard example, m/s is smaller than diameters of elementary particles relative to macroscopic terrestrial objects.

We now want to show that the intensity of capital on the envelope is, except at rates of profit very close to zero or to R_{σ_s}, very close to the intensity of capital of w^6, which is very close to the intensity of capital of the diagonal d, interpreted as a potential wage curve with intensity of capital $w^1(0)/R_{\sigma_s}$ (equal to one with our normalization). Indeed, it is geometrically obvious that $\frac{w^6(0)}{R_6} \cong \frac{w^1(0)}{R_{\sigma_s}}$ since $w^6(0)$ is in I_w and R_6 is in I_R. If w^6 is not part of the envelope, the envelope itself is even closer to d.

We can summarize broadly as follows: Having fixed γ in order to reach a desired degree of certainty, we can still increase s, if the technology is rich enough in methods, and in theory we can go with s to infinity so as to bring the envelope, which consists of line segments, arbitrarily close to d, so that the capital-intensities, represented by the line segments, must approach the slope of d, since the envelope as a whole is monotonically falling.

Proposition 3
The intensity of capital tends, except in a vanishingly small neighbourhood of $r = 0$ and $r = R_s$, to $w^1(0)/R_{\sigma_s}$, as s tends to infinity as in Corollary 3.

Matters look different for finite s, as one approaches the corners E and B. The envelope may start at E with a nearly vertical slope (the intensity of capital is large initially) and ends at B

with a tangentially horizontal slope (the intensity of capital may be small). It then seems to follow from our assumptions that the envelope has the properties of the wage curve one derives from a neoclassical production function, but only in the small, at small rates of profit or wage rates. In a world with a large, but finite s, the envelope could therefore look as in Diagram 3, consisting of a large number of line stretches.

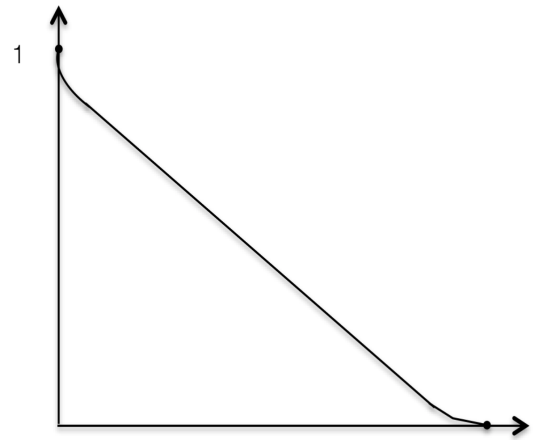


Diagram 3: The envelope for many techniques .

The absolute value of the slope of the first wage curve in going down from E and of the last, when one arrives at B, is a priori anywhere between infinity and the slope of the diagonal, and between that and zero respectively, so that these initial and terminal wage curves determine the slopes of the beginning and the end of the envelope. The probability that the first wage curve is the steepest is the same as that it is already the dominating wage curve and actually equal to d, namely $1/s$. Moreover, as s increases, point P_6 will be driven into the upper right corner, as m/s tends to zero. This means that the wage curve corresponding to P_6, w^6, will tend to the diagonal also near the axes, and the possible wage curves of high and low capital-intensity there will tend to get dominated so that the intensity of capital will, in our normalization, tend to one near both ends. There remains little room for the substitution of capital for labour with these assumptions, and Joan Robinson's position then is vindicated.

To assume a smooth wage curve as in Diagram 3 is a concession to neoclassical theory. Our original assumptions define an envelope composed of line segments of a subset of the wage curves $w^\sigma(r)$. As one moves down the envelope, the wage rate falls monotonically as the rate of profit rises. Each line segment of a $w^\sigma(r)$ on the envelope is associated with a definite capital intensity k_σ, which falls as r rises, and an output per head $y_\sigma = w^\sigma(0)$, which also falls. If one associates each of the capital intensities, going now from the lowest to the highest, with the corresponding output per head, $y_\sigma = w^\sigma(0)$, one obtains a rising step function $y_\sigma = f(k_\sigma)$, and the question is whether a smoothing of the envelope will lead to a smooth production function fulfilling the marginal productivity conditions.

We now postulate that a smooth envelope $w(r)$ can be given with stylized assumptions, that is with $w'(0) = k_{\sigma_s} \gg 1$, $w'(R_{\sigma_s})$ small and, as in Diagram 3, with a long linear stretch in the middle between rates of profit r_1 and r_2, as an expression of the fact that the envelope approximates the diagonal of the wage curve box, like wage curve $w^6(r)$ in Diagram 2 according to what we have proved. The intensity of capital k is equal to the absolute value of the slope of $w(r)$ at each r and $y = f(k)$ is given by the intersection of the tangent with the ordinate so that $f(k)$ rises monotonically and continuously. The intensity of capital is equal to one along the linear segment of $w(r)$ between r_1 and r_2, if we assume a normalised wage curve box, and $f(1)$ is well defined, but $f(k)$ has a kink at $k = 1$, since the rate of profit equals r_2, if we approach $k = 1$ from below, and r falls from r_1, as k is raised beyond $k = 1$. This indeterminacy of the marginal productivity condition is not small, since r_1 tends to zero and r_2 to R_{σ_s}, as s increases.

To show this more formally, we have to go back to the well-known mathematical construction of the production function, starting from the wage curve (Samuelson 1962, p. 202; Schefold 1989, pp. 297-8; the procedure is the reversal of the derivation of the wage curve from the production function explained in Note 7).

For each individual technique we have output per head $y_\sigma = w^\sigma(0)$ because of the choice of the numéraire (Note 2). Accounting yields $(y_\sigma - w(r))/r = k$, and the tangent to the wage curve $w(r)$ at r has slope $-k_\sigma$, because $w(r)$ is the envelope of the wage curves. Hence $k = -w'(r)$ at each r, and with this the production function can be defined parametrically by $f(k) = w(r)+rk = w(r) - rw'(r)$. Now $f(k)$ is well defined, if $w''(r) \neq 0$, which is the case in the intervals $[0, r_1)$ and in $(r_2, R_{\sigma_s}]$, but $w''(r) = 0$ in $[r_1, r_2]$. We have in $[0, r_1)$ and in $(r_2, R_{\sigma_s}]$ as in Note 7

$$\frac{df}{dk} = \frac{df}{dr}\Big/\frac{dk}{dr} = \frac{1}{-w''}(w' - w' - rw'') = r.$$

This means that the production function, so defined, fulfils the first-order condition of the marginal productivity relationship (and similarly for the second order), but the linear section of $w(r)$, where the function $k = w'(r)$ cannot be inverted, results in a kink of f at $k = 1$.

Proposition 4
Given the stylized assumptions of Diagram 3, an approximate surrogate production function $f(k)$ results, which fulfils the marginal productivity condition $f'(k) = r$ and $f'' < 0$, except at $k = 1$, where the function is continuous, but the left-hand derivative equals r_2 and the right-hand derivative equals r_1.

The indeterminacy of the marginal product is not small, for we get from Proposition 3:

Corollary
As s tends to infinity, r_2 tends to $R_{\sigma_s} = 1$ and r_1 to zero.

Since the economy will tend to a state with $k = 1$, distribution remains unexplained. The underlying reason will emerge in the next section, where it is shown that the economy does not only tend to a specific value of the capital-labour ratio, but that the number of techniques that are actually eligible is surprisingly small. This means that, if we do not flatten the envelope by introducing a continuum of techniques and by assuming, à la Samuelson, that it is as smooth as in Diagram 3, there will be substitution in the sense of changing of techniques at the (small number of) switchpoints, but there will be virtually no substitution in the (from the point of view of neoclassical theory) relevant sense of changes of the capital-labour ratios, for they tend to equality, except at the ends of the envelope, as will be shown more rigorously in the next section. This, then, is the new critique of neoclassical theory, which results on the basis of the Samuelsonian assumption that wage curves are linear, irrespective of whether one regards this assumption as a concession in the debate or because one regards it as licit in the belief that empirical wage curves are quasi-linear or because one argues that wage curves must tend to be quasi-linear as a result of the random character of systems. For extensions to the non-linear case, see Section 5, for empirical confirmation see Section 6.

4 Theorem by Kersting

The theorem is introduced as a mathematical problem with its own notation. It may be studied independently of the other sections of the paper, but it provides an essential insight for the economic analysis, by which it is inspired.

Let $s \in \mathbb{N}$ and let $(\sigma_1, \ldots, \sigma_s)$ be a random permutation of $(1, \ldots, s)$. Denote by w^k the straight line in the plane passing through the points k at the ordinate and σ_k at the abscissa. We are going to study their envelope w in the plane's first quadrant,

$$w(r) := \max_{1 \le k \le s} w^k(r), \quad r \in [0, s].$$

Between its endpoints $(0, s)$ and $(s, 0)$ it is made up of several line segments. Let K be the set of

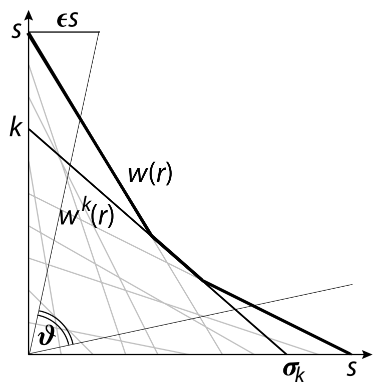


Diagram 4: The envelope with $s = 10$.

points in the plane, where these segments join up (the "kinks" of w), and denote its cardinality by

$$X_s := \#K.$$

Also, for $0 < \epsilon < 1$ set

$$X_{s,\epsilon} := \#\{(i, j) \in K : \epsilon i \le j \text{ and } \epsilon j \le i\},$$

which is the number of kinks in the sector within the first quadrant with the angle $\vartheta = \pi/2 - 2 \arctan \epsilon$. Note that $X_s = X_{s,1/s}$.

Theorem. *As $s \to \infty$*

$$E[X_s] \sim \frac{2}{3} \ln s \quad \textit{and} \quad E[X_{s,\epsilon}] \sim \frac{2}{3} \ln \epsilon^{-1}.$$

The intuition behind this result is as follows: The envelope w is primarily made up of a few lines w^k, which have the property that both k and σ_k are close to s (the deviation being of order $\sqrt{s}$). They show up in the expectations $E[X_{s,\epsilon}]$, with increasing s their number remains bounded.

Additionally, w contains at both ends lines w^k where just k or just σ_k is close to s, including the lines with $k = s$ or $\sigma_k = s$. Our result indicates that there are on average $\frac{1}{3} \ln s$ many at each end.

The same result appears, if the terms $w^k(0)$ and σ_k do not follow a regular pattern, but arise as values of independent random variables, uniformly distributed on the interval $(0, 1)$. Since the envelope is primarily made up from those terms $w^k(0)$ and σ_k close to 1, the requirement of uniformity may well be relaxed. It is enough to assume that the independent random variables have a continuous density f on the interval $(0, 1)$ fulfilling $f(1) > 0$. It is a challenging question, to what extent the situation changes for densities with the property $f(x) > 0$ for all $x > 0$. This problem will be treated in a future study.

The next figure illustrates these asymptotic results by simulations from samples of size 10.000. The dots in the left-hand illustration specify the simulated expectations $E[X_s]$ with $s = 10^a$, $a = 1, \dots, 9$, and the line is the function $\frac{2}{3} \ln s$. The dots in the right-hand graphics give the simulated expectations of $E[X_{s,\epsilon}]$ for the two values of ϵ when $\frac{2}{3} \ln \epsilon^{-1}$ is equal to 1 or 2. They correspond to the angles ϑ with degrees 64.8 and 84.3, respectively.

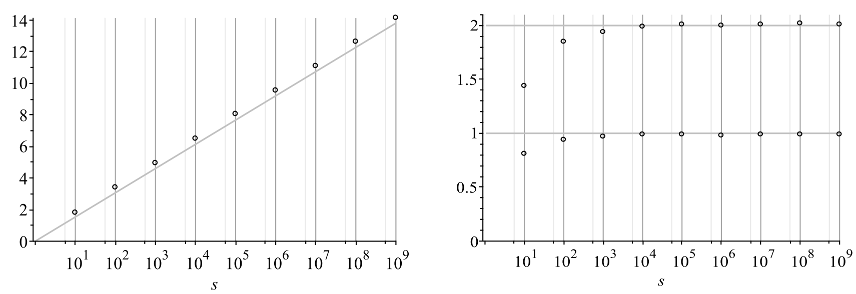


Diagram 5: The simulations.

(Outline of) Proof. The full proof will be given in a separate paper. Here we omit some technical details and focus on the central issues.

(i) Let us call a quadrupel (a, b, c, d) of natural numbers a *constellation*, if $1 \le a < b \le s$ and $1 \le c < d \le s$. It determines the point (i, j) of intersection of the straight line between

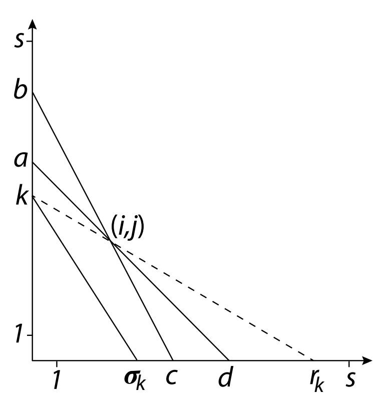


Diagram 6: A constellation.

$(0,a)$ and $(d,0)$ and the line between $(0,b)$ and $(c,0)$. Each element of K is such an intersection point, which suggests to count K by deciding for each intersection point, whether it belongs to K or not. Thus, we assign to each constellation (a,b,c,d) a random variable, indicating that it participates in the envelope in the sense that both lines from a to d and from b to c take part in w. In other terms:

$$Y_{abcd} := \begin{cases} 1, & \text{if } \sigma_a = d, \sigma_b = c \text{ and } (i,j) \in K, \\ 0 & \text{else.} \end{cases}$$

Each element of K will be captured by one of the random variables. Some element of K may be multiply recorded, since we have $(i,j) = (i',j')$ for some pairs of constellations, however, this feature is negligible in the limit $s \to \infty$. Summing these random variables over the constellations fulfilling $\varepsilon i \leq j \leq i/\varepsilon$ yields

$$X_{s,\epsilon} \leq \sum_{\substack{a,b,c,d \\ \epsilon i \leq j \leq i/\epsilon}} Y_{abcd},$$

and the expectation taken on the left-hand side is well approximated by that on the right, that is

$$E[X_{s,\epsilon}] \sim \sum_{\substack{a,b,c,d \\ \epsilon \leq j/i \leq 1/\epsilon}} P(Y_{a,b,c,d} = 1) \tag{1}$$

as $s \to \infty$. To a certain extent this approximation is valid also for variable ε. In particular, inserting $\varepsilon = 1/s$ we obtain an approximation for the expectation of $X_s = X_{s,1/s}$.

(ii) Next we deduce a formula for the probability of $\{Y_{abcd} = 1\}$. The occurence of this event requires that none of the lines w^k runs above the point (i,j). This is immediate for $k \leq j$, and for $k > j$ it is required that $w^k(i) \leq j$. Since $w^k(i) = k(1 - i/\sigma_k)$, the latter condition boils down to

$$\sigma_k \leq r_k \quad \text{with} \quad r_k := \frac{ik}{k-j}.$$

The requirement has to be taken into account as long as $r_k < s$ or, equivalently, as long as

$$k > \frac{js}{s-i}. \tag{2}$$

It follows that

$$\{Y_{abcd} = 1\} = \{\sigma_a = d, \sigma_b = c \text{ and } \sigma_k \leq r_k \text{ for all } \frac{js}{s-i} < k \leq s\}.$$

Also observe that r_k is increasing with decreasing k. Consequently, placing the lines w^s, w^{s-1}, ... one after the other, and checking all favorable and possible outcomes, we obtain the formula

$$\begin{aligned} P(Y_{abcd} = 1) = \frac{[r_s]}{s} \frac{[r_{s-1}] - 1}{s-1} \cdots \frac{1}{b} \cdots \frac{1}{a} \cdots &= \frac{1}{ab} \prod_{\substack{\frac{js}{s-i} < k \leq s \\ k \neq a,b}} \frac{[r_k] - (s-k)}{s-(s-k)} \\ &= \frac{1}{ab} \prod_{\substack{\frac{js}{s-i} < k \leq s \\ k \neq a,b}} \Big(1 - \frac{s - [r_k]}{k}\Big) \end{aligned}$$

where $[r_k]$ denotes the biggest natural number not exceeding r_k. Incorporating in addition the inequalities $1-x \leq e^{-x}$ and $[r_k] \leq r_k$ yields

$$P(Y_{abcd}=1) \leq \frac{1}{ab} \exp\Big(- \sum_{\substack{\frac{js}{s-i}<k\leq s \\ k\neq a,b}} \frac{s-r_k}{k} \Big). \tag{3}$$

It can be shown that in our calculations we may substitute this upper bound for the probability.

(iii) Now we derive an approximation for the right-hand expression in (3). From a heuristic point of view it is obvious that, with s increasing, there are lines w^k with the property that both k and σ_k don't deviate much from s. This has been shown more formally in Section 3 (Proposition 2 and Corollaries). This implies that the envelope w is close to the diagonal connecting $(0,s)$ and $(s,0)$. Therefore it is plausible, and indeed can be shown, that we may confine our considerations to constellations, fulfilling

$$a,b,c,d \sim s$$

as $s \to \infty$. In order to use these asymptotics for i and j we introduce the notations

$$x := b-a\ ,\ y := d-c\ ,\ s-b := u\ ,\ s-d := v.$$

A quick calculation results in

$$i = \frac{cd(b-a)}{c(b-a)+b(d-c)} \sim \frac{sx}{x+y}\ ,\ j = \frac{ab(d-c)}{c(b-a)+b(d-c)} \sim \frac{sy}{x+y}, \tag{4}$$

hence

$$s-i \sim \frac{sy}{x+y} \sim j\ ,\ s-j \sim \frac{sx}{x+y} \sim i,$$

moreover

$$s-i-j = \frac{c(b-a)(s-d)+b(d-c)(s-a)}{c(b-a)+b(d-c)} \sim \frac{xv+yu+xy}{x+y}.$$

Coming back to (3) observe that for $k \leq s$ fulfilling (2) we have $k \sim s$. Consequently

$$\frac{s-r_k}{k} = \frac{s-i}{k(k-j)}\Big(k - \frac{js}{s-i}\Big) \sim \frac{s-i}{s(s-j)}\Big(k - \frac{js}{s-i}\Big) \sim \frac{y}{sx}\Big(k - \frac{js}{s-i}\Big)$$

and

$$\begin{aligned} \sum_{\frac{js}{s-i}<k\leq s} \frac{s-r_k}{k} &\sim \frac{y}{sx} \sum_{\frac{js}{s-i}<k\leq s} \Big(k - \frac{js}{s-i}\Big) \\ &\sim \frac{y}{sx} \int_{\frac{js}{s-i}}^{s} \Big(z - \frac{js}{s-i}\Big)\, dz \\ &= \frac{y}{sx}\frac{1}{2}\Big(s - \frac{js}{s-i}\Big)^2 \\ &= \frac{sy(s-i-j)^2}{2x(s-i)^2} \\ &\sim \frac{f(x,y,u,v)}{2s} \end{aligned}$$

with

$$f(x,y,u,v) := \frac{(xv+yu+xy)^2}{xy}, \; x,y \ge 1 \;, \; u,v \ge 0.$$

Applying the approximations to (3) (and neglecting the requirement $k \neq a,b$ for the summation index) yields

$$P(Y_{abcd}=1) \sim \frac{1}{s^2} \exp\Big(-\frac{f(x,y,u,v)}{2s}\Big).$$

To arrive at the expectation these terms have to be summed for all constellations (a,b,c,d), that is for all natural numbers $x,y \ge 1$ and $u,v \ge 0$ with $x+u<s, y+v<s$. Again it can be shown that the latter requirements can be ignored in the limit $s \to \infty$. Also in view of (4) we replace the condition $\varepsilon \le j/i \le 1/\varepsilon$ by $\varepsilon \le y/x \le 1/\varepsilon$. With these adjustments (1) yields

$$E[X_{s,\varepsilon}] \sim \frac{1}{s^2} \sum_{\substack{x,y\in\mathbb{N},u,v\in\mathbb{N}_0 \\ \varepsilon \le y/x \le 1/\varepsilon}} \exp\Big(-\frac{f(x,y,u,v)}{2s}\Big)$$

and also

$$E[X_{s,\varepsilon}] \sim \frac{1}{s^2} \iiiint\limits_{\substack{x,y,u,v>0 \\ \varepsilon \le y/x \le 1/\varepsilon}} \exp\Big(-\frac{f(x,y,u,v)}{2s}\Big)\, dxdydudv \;. \tag{5}$$

(iv) It remains to determine this fourfold integral. First we accomplish the integration with respect to u and v for fixed x and y. To this end we introduce new coordinates

$$z := \frac{xv+yu+xy}{\sqrt{xy}} \;, \; w := xu - yv \;.$$

The condition $u,v \ge 0$ translates into

$$z \ge \sqrt{xy} \quad , \quad -\frac{y}{x}(z\sqrt{xy}-xy) \le w \le \frac{x}{y}(z\sqrt{xy}-xy) \;,$$

and for the Jacobian determinant we have $|\frac{\partial(z,w)}{\partial(u,v)}| = -(x^2+y^2)/\sqrt{xy}$, thus $dudv = \sqrt{xy}\, dzdw/(x^2+y^2)$. Hence

$$\begin{aligned}
&\iint\limits_{u,v>0} \exp\Big(-\frac{(xv+yu+xy)^2}{2sxy}\Big)\, dudv \\
&\qquad = \int_{\sqrt{xy}}^{\infty} \int_{-\frac{y}{x}(z\sqrt{xy}-xy)}^{\frac{x}{y}(z\sqrt{xy}-xy)} e^{-z^2/(2s)} \, \frac{\sqrt{xy}\, dwdz}{x^2+y^2} \\
&\qquad = \int_{\sqrt{xy}}^{\infty} \Big(\frac{x}{y}+\frac{y}{x}\Big)(z\sqrt{xy}-xy)e^{-z^2/(2s)} \frac{\sqrt{xy}}{x^2+y^2}\, dz \\
&\qquad = \int_{\sqrt{xy}}^{\infty} (z-\sqrt{xy})e^{-z^2/(2s)}\, dz \;.
\end{aligned}$$

Next we complete the calculation by performing the intergration with respect to x and y. Here a second change of coordinates is helpful. We set $\eta = xy$, $\xi = y/x$ with corresponding Jacobian determinant $|\frac{\partial(\eta,\xi)}{\partial(x,y)}| = 2y/x = 2\xi$. Thus $dxdy = d\eta d\xi/2\xi$ and

$$\iiiint\limits_{\substack{x,y,u,v>0\\ \varepsilon\le y/x\le 1/\varepsilon}} \exp\Big(-\frac{(xv+yu+xy)^2}{2sxy}\Big)\,dxdydudv$$

$$\begin{aligned}
&= \iint\limits_{\substack{x,y>0\\ \varepsilon\le y/x\le 1/\varepsilon}} \int_{\sqrt{xy}}^{\infty}(z-\sqrt{xy})e^{-z^2/(2s)}\,dz\,dxdy \\
&= \int_{\varepsilon}^{1/\varepsilon}\int_0^{\infty}\int_{\sqrt{\eta}}^{\infty}(z-\sqrt{\eta})e^{-z^2/(2s)}\,dz\frac{d\eta d\xi}{2\xi} \\
&= \int_{\varepsilon}^{1/\varepsilon}\frac{d\xi}{2\xi}\int_0^{\infty} e^{-z^2/(2s)}\int_0^{z^2}(z-\sqrt{\eta})\,d\eta dz \\
&= \ln\varepsilon^{-1}\int_0^{\infty}\frac{1}{3}z^3e^{-z^2/(2s)}\,dz.
\end{aligned}$$

By means of the subsitution $z' = z^2/(2s)$ it follows that the right-hand integral has the value $2s^2/3$. Going back to (5) we obtain

$$E[X_{s,\varepsilon}] \sim \frac{2}{3}\ln\varepsilon^{-1},$$

which is one of our claims. It holds also true, if we replace ε by $1/s$, which gives the other claim. $\square$

5 Reconsidering results and assumptions

The unusual findings at which we have arrived will provoke doubts and discussions. We shall try to respond to some potential objections. Before doing so, we synthesize some earlier results in a concrete economic example. Most economists do not take the production function for an immediate image of reality, but as a construction, which helps to visualize some aspects of what they perceive to be the basic characteristics of the economy: the production function illustrates the substitution of capital for labour, as labour becomes scarce and capital is accumulated in the long run or the substitution of labour for capital, if there is unemployment. In either case, substitution means that capital goods are recombined in different techniques – new ones in the process of growth, old ones, if unemployment is to be absorbed. In the former case, labour stays roughly constant and capital grows, *vice versa* in the latter. This happens in the relevant range of the rate of profit, and profit maximisation implies that the techniques chosen are on the envelope and efficient.

Diagram 7 shows how the substitution possibilities present themselves in our model. Suppose the industrial rate of profit is between r_1, say 3%, and r_2, say 20%, and the assumptions of the theorem of Section 4 hold. According to the earlier propositions, the envelope will be so close to the diagonal, that the distance between them may be neglected. This made it possible to conclude that the capital-labour ratio cannot change much on the envelope, except at its ends. Thanks to the Theorem, we can now turn to the change of techniques itself and calculate the number of substitutions (switch-points) that can be expected between r_1 and r_2 by calculating the difference between the numbers of switch-points of the larger cone with angle ϑ_1, defined by r_1, and the narrower cone with angle ϑ_2, defined by r_2, in Diagram 7. For reasons of symmetry, one half of the difference must be taken.

The corresponding ε_1 and ε_2 result from the elementary intercept theorem as $\varepsilon_i = r_i/(1-r_i)$, assuming that we have the square wage curve box with side length 1 and a regular grid. The expected number of switch-points between r_1 and r_2 then is

$$\frac{1}{2}\Big(\frac{2}{3}\ln\frac{1}{\varepsilon_1} - \frac{2}{3}\ln\frac{1}{\varepsilon_2}\Big) = \frac{1}{3}\Big(\ln\Big(\frac{1}{r_1} - 1\Big) - \ln\Big(\frac{1}{r_2} - 1\Big)\Big).$$

With $r_1 = 3\%$ and $r_2 = 20\%$ one gets 2.08/3. Hence less than one switch-point, hence only one technique is expected to be encountered in this fairly large relevant range of the rate of profit. The capital-intensities of neighbouring techniques will for large s nearly be the same in a much larger range, but here, there is not even a method change, and this despite the fact that, the number of techniques between zero and r_1 will tend to infinity as $s \to \infty$. The result is perplexing: as more techniques become available, fewer will appear in the relevant range, since most switch-points are pushed into the corners. The economy will stay at $k = 1$, if we represent it by means of an approximate surrogate production function (Proposition 4), and it is confirmed that output per head will be near one, as long as the rate of profit remains in the relevant range.

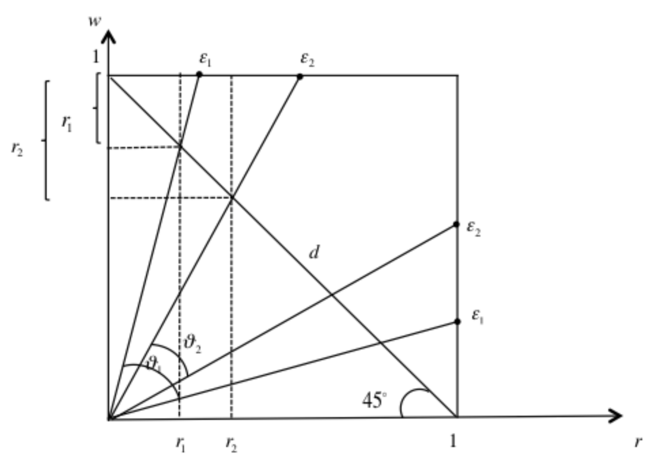


Diagram 7: Two cones (angles ϑ_1 and ϑ_2) defined by r_1 and r_2. The envelope within the difference of the cones between r_1 and r_2 is likely to contain only one technique with $r_1 = 3\%$ and $r_2 = 20\%$.

We concluded at the end of Section 3 that there is virtually no substitution between capital and labour, if efficient techniques are considered, and we now see that there is, in the relevant range, then hardly any method change at all. We mention important conclusions before we discuss the assumptions.

An important difference between this critique of the neoclassical approach and the older one, based on reswitching and reverse capital deepening, should be noted. Reswitching and reverse capital deepening are rare, as we saw by means of references to earlier work (Section 1) both empirically and, for large systems, in theory. Almost universally dominating techniques also are rare: rare in the set of all techniques. But while it does not matter much – it probably will not even get noticed – when reverse capital deepening occurs in reality, almost universally dominating techniques are reached or approximated through profit maximisation and show in the fact that actual techniques change through progress, not distribution.

Where did the idea of substitution as the central mechanism regulating distribution come from? The idea of a process of accumulation accompanied by a rising capital-labour ratio was, as we just have recalled, suggested by the rise of machinery in Ricardian times (Schefold, 1976). Since there was labour-saving technical progress, one could also conceive of the opposite: to opt for less capital-intensive techniques to employ more labour, and factor prices would steer these processes. The principle of the process of substitution had been well understood in the case of land and labour thanks to the Ricardian theory of intensive rent. But capital is not land. It was found that the gain in labour productivity lowered not only the cost of output in real terms, but also that of capital goods so that the capital-output ratio could stay constant, which is the inverse of the maximum rate of profit. Accumulation could proceed at a constant rate of profit $\overline{r}$ and with constant shares of profits and wages, which means formally that the wage curve turns around the point of the maximum rate of profit $r = R$. Diagram 8 illustrates these 'stylized facts'.[21] The wage curves shown are those of the universally dominating techniques of each period. A change in the rate of profit – whatever causes the change of distribution – does not lead to a substitution with significant changes of the capital-labour ratio. The latter changes only in the long run through technical progress.

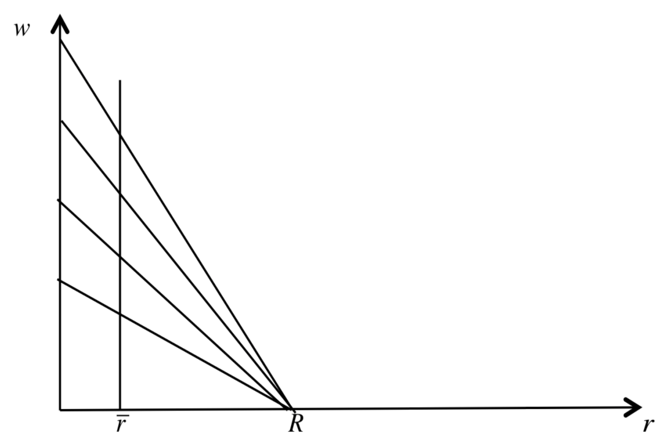


Diagram 8: A growing economy; the dominating technique of each period is represented by a wage curve showing the growth of the capital-labour ratio at a maximum rate of profit (inverse of the capital-output ratio) that stays constant.

One could summarize this last result by stating that our approach provides a theory of capital that can serve as the foundation for a Kaldorian theory of growth. This means that the new turn in capital theory has a positive consequence, by which it distinguishes itself from the negative critique based on reswitching and related arguments. The old critique could not predict how the capital-labour ratio would change with changes in the rate of profit, while it emphasized that any change of the rate of profit could trigger an avalanche of changes of methods of production. This critique was agnostic as to what the method change would imply for employment or the ratio of profits to wages. The new critique leads back to the old classical method of regarding the actual technique in existence as a given; it changes only slowly. The system tends towards an efficient technique with a stable capital-labour ratio. Thus, room is made for alternative theories of distribution such as a postkeynesian determination of the share and the rate of profit via demand or the monetary theory of distribution, which explains the level of the rate of profit through the intermediate influence of the rate of interest. I argue in Schefold (2021) that these alternative theories of distribution presuppose a given technique and are not really compatible with the old critique of capital and its agnosticism.

But do the assumptions hold? We here try a theoretical discussion of alternative assumptions, which is unconventional because it does not start from usual assumptions about distributions. We must go beyond the strong case. We begin by reconsidering the central hypothesis itself. Readers not interested in this experiment may go to Section 6 directly and take note only of the subsection on Non-linear Wage Curves. We limit ourselves to the discussion of Proposition 1 and the domination from above. There is a large number s of techniques with by assumption $w^1(0) > \cdots > w^s(0)$. We found $P(R_\sigma > R_{\sigma-1}, \ldots, R_\sigma > R_1) = 1/\sigma$; this holds, if we put $R_0 = 0$, for $\sigma = 1, \ldots, s$. Now we ask whether our conclusions are affected substantially, if we assume that the die is loaded somehow in favour of or against the neoclassical assumption, in that a lowering of σ – the transition from $w^{\sigma-1}(0)$ to $w^\sigma(0)$ – tends to lead to $R_\sigma > R_{\sigma-1}$ with a probability that is higher or lower than in the strong case. Countless variations of this idea can be imagined. We shall depart from the strong case in Section 6 by introducing bivariate distributions, which are not uniform, with different correlations. Here, we consider instead a simple possibility for modelling different pairings of the $w^\sigma(0)$ and the R_σ more directly, without specifying assumptions about the distributions of $w^\sigma(0)$ and R_σ on the axes. Only one example for an alternative distribution will be constructed.

If we put $P(R_\sigma > R_{\sigma-1}, \ldots, R_\sigma > R_1) = \frac{\lambda}{\sigma}$, $\sigma = 2, \ldots, s$, we get for the estimate from above for the expected number of wage curves on the envelope, ω, now $\omega \cong \lambda \ln s$: no significant insight and no improvement from the point of view of neoclassical theory.

An interesting possibility is to postulate $P(R_\sigma > R_{\sigma-1}, \ldots, R_\sigma > R_1) = \frac{1}{\sigma^\beta}$, $\beta > 0$, $\beta \neq 1$; $\sigma = 1, \ldots, s$. The case $\beta = 1$ is what we have been discussing in Section 3. If $\beta \neq 1$, one obtains

$$\omega \cong \int_1^s \frac{1}{x^\beta}\, dx = \frac{x^{1-\beta}}{1-\beta}\bigg|_1^s.$$

Only for completeness we mention that one gets $P(R_\sigma > R_{\sigma-1}, \ldots, R_\sigma > R_1) < 1/\sigma$ for $\beta > 1$ and, for large s, $\omega \cong 1/(\beta - 1)$, for instance with $\beta = 3/2$, $\omega \cong 2$, which means that there is virtually no substitution to be expected.

Now, what about the case $\beta = 0$? If all techniques are to appear on the envelope with certainty, the order of the maximum rates of profit must obviously be exactly inverse to that of output per head; one then must have that $w^1(0) > \cdots > w^s(0)$ implies $R_1 < \cdots < R_s$. These relationships must hold for the wage curves on the envelope as the result of the choice of efficient techniques. These must be such that any lowering of the productivity of labour results in an increase of the productivity of capital. To assume these relationships instead of deriving them would not mean to prove the existence of the neoclassical production function; it means simply to postulate it.

If $0 < \beta < 1$, $\omega \cong \big(1/(1-\beta)\big)(s^{1-\beta} - 1)$, hence, for $\beta = 1/2$, $\omega \cong 2(\sqrt{s} - 1)$. We now do get more techniques on the envelope, but still $\Omega = \omega/s \to 0$ as $s \to \infty$; the set of techniques appearing on the envelope remains infinitesimal relative to the set of all techniques. What is more important: the possibility that a $w^{\sigma-1}$ associated with a small $R_{\sigma-1}$ is followed by a w^σ with a R_σ close to R_{σ_s} is not only not excluded by the assumption, but it becomes increasingly probable with rising s even for some small σ. In other words, there will in this case as well appear anchor points in the upper right corner of the wage curve box in Diagram 2, only more slowly, and this means that almost universally dominant techniques will exist in this case as well and the envelope must approach the diagonal d with very large s. The proof requires an assumption, however, as we shall show (Proposition 5).

We analyse the case of $0 < \beta < 1$ by means of a representative example, putting $\beta = 1/2$. We begin by constructing a series of anchor points, which correspond to an envelope with $\beta = 1/2$. Let $Q(\sigma)$ be the number of positive square integers smaller or equal to σ and $\lambda = 1, 2, 3, ...,$ $\sigma_\lambda = \lambda^2 = 1, 4, 9, ...$ Hence $Q(\sigma_\lambda) = Q(\lambda^2)$ is the number of elements of the set $\{1, 4, 9, ..., \lambda^2\}$, which is equal to λ; $Q(\sigma_\lambda) = \lambda = \sqrt{\sigma_\lambda}$. If we have an envelope E with the property

$$\sigma = \begin{cases} \lambda^2 & \text{, then } R_\sigma > R_{\sigma-1}, ..., R_\sigma > R_1 \\ \text{otherwise} & \text{, then } \exists R_\tau, \tau < \sigma : R_\tau > R_\sigma, \end{cases}$$

it implies for E

$$P(R_\sigma > R_{\sigma-1}, ..., R_\sigma > R_1) \cong \frac{1}{\sqrt{\sigma}}$$

$$\text{for } P = \frac{Q(\sigma)}{\sigma} = \frac{\sqrt{\sigma}}{\sigma} = \frac{1}{\sqrt{\sigma}}.$$

Such an envelope can be constructed, using the following two rules:

(i) If $\sigma = \sigma_\lambda = \lambda^2$, then define $R_{\sigma_\lambda} = (\lambda + 1)^2 - 1 = \lambda^2 + 2\lambda$.

The rule implies $\sigma_{\lambda-1} = (\lambda - 1)^2$ and $R_{\sigma_{\lambda-1}} = \lambda^2 - 1$. Hence there is a sequence of 2λ $w^\sigma(0)$ in between $\sigma_{\lambda-1}$ and σ_λ and there are 2λ free positions between $R_{\sigma_{\lambda-1}}$ and R_{σ_λ}.

(ii) The maximum rates of profit between $R_{\sigma_{\lambda-1}}$ and R_{σ_λ} can be assigned according to any permutation to the $w^\sigma(0)$ in the interval between $\sigma_{\lambda-1}$ and σ_λ, but not to any $w^\sigma(0)$ outside this interval.

If this last condition is not fulfilled, additional switch points may arise, destroying the property $P = \frac{1}{\sqrt{\sigma}}$. Diagram 9 shows the anchor points in the wage curve box for this envelope up to $\lambda = 4$ and with the simple specification for rule (ii) $R_\sigma = \sigma - 1, (\sigma \neq \lambda^2)$.

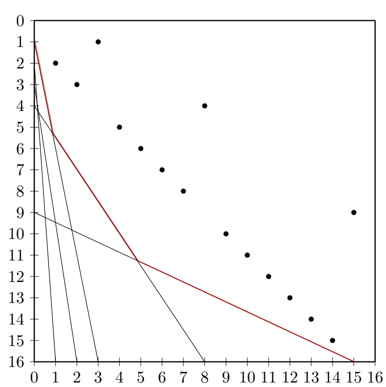


Diagram 9: Wage curve box with anchor points for the case $\beta = 1/2$. The first three segments of the envelope have been drawn and the intersecting first two wage curves that remain below the envelope. The outer sequence of the anchor points is at distance 2λ from the diagonal and not linear.

The example suggests that, as s increases in a wage curve box of given size and with $0 < \beta < 1$, the envelope will look more and more like Samuelson's 'hyperbolic' 'smooth' wage curve that could give rise to a smooth production function with diminishing returns. However, this conclusion holds only, if the probability of finding 'good' techniques, represented by anchor points in the upper right corner of the wage curve box, is very small. This is not plausible and, as we shall see in the next section, against the evidence, if the wage curve box is bounded. The envelope of Diagram 9 is based on a discrete distribution in a regular grid. The distribution of the $w^\sigma(0)$ and the R_σ on the axes is uniform. However, their pairings are not independent, but such that $P(R_\sigma > R_{\sigma-1}, ..., R_\sigma > R_1) = \frac{1}{\sqrt{\sigma}}$. We therefore abandon the assumption that there is one and only one anchor point on each line and column. Instead, there is on each line σ one anchor point in position τ with a probability σ_τ; $0 \leq \sigma_\tau \leq 1$; $\sigma_1 + ... + \sigma_s = 1$. The distribution of the $w^\sigma(0)$ and the R_σ on the axes remains uniform. There is no room here to analyse the distribution of the σ_τ in order to derive the conditions under which $P(R_\sigma > R_{\sigma-1}, ..., R_\sigma > R_1) = 1/\sigma^\beta$ results. It must suffice that Diagram 9 exemplifies the possibility for a determinate matrix of the σ_τ, $\sigma_\tau = 1$ or $\sigma_\tau = 0$, according to the rules (i) and (ii). If the probabilities σ_τ are themselves distributed uniformly, given the uniform distributions of the $w^\sigma(0)$ and the R_σ, one is essentially lead back to the strong case of Sections 3 and 4. But here we have modified the distribution of the σ_τ and prove the following generalised variant of Proposition 2 and its Corollaries.

Proposition 5

Let a bivariate probability distribution for the anchor points in the wage curve box be given, such that the probabilities for finding an anchor point in each row or column add up to one and such that all probabilities in the triangle DFG (Diagram 2) are greater or equal to $\delta^\star = \delta/s$, with $\delta > 0$ and with m, the length of the short sides of the triangle given by $m = \sqrt{\gamma s}$, where $1 \leq \gamma < s$. Then there exists, with at least probability $\pi^\star = 1 - 1/\sqrt{e^{\gamma\delta}}$, a wage curve $w^\star$ in the wage curve box with a distance from the diagonal (measured horizontally or vertically) not larger than m.

Even if δ is small, the choice of γ can bring $\pi^\star$ close to one, and m will remain small relative to large s, for $m/s = \sqrt{\gamma/s}$.

The proof still refers to Diagram 2. In order to show that the complementary probability of $\pi^\star$ tends to zero, observe that the probability for an anchor point to be on the first line of the trapezium $CDGE$ is at most $\mu_1 = 1 - m\delta^\star$, on the second line $\mu_2 = 1 - (m-1)\delta^\star$, etc. On the last line it is $\mu_m = 1 - \delta^\star$. The probability for all anchor points to be in $ABDGE$ is $\mu_1 \cdot ... \cdot \mu_s = \mu_1 \cdot ... \cdot \mu_m$, since anchor points are on the remaining rows with probabilities $\mu_\sigma = 1$. Now, putting $m = \sqrt{\gamma s}$ and using the same approximations as in the proof of Proposition 2,

$$\begin{aligned}
\mu_1 \cdot ... \cdot \mu_m &= \exp(\ln[(1 - m\delta^\star) \cdot ... \cdot (1 - \delta^\star)]) \\
&\cong [\exp(m + (m-1) + ... + 1)\delta^\star]^{-1} \\
&= [\exp \frac{m(m+1)}{2}\delta^\star]^{-1} \cong [\exp \frac{\delta}{2s}(\gamma s + \sqrt{\gamma s})]^{-1} \\
&\xrightarrow[s \to \infty]{} e^{-\frac{\delta\gamma}{2}},
\end{aligned}$$

hence we have at least

$$\pi^\star = 1 - \frac{1}{\sqrt{e^{\delta\gamma}}}.$$

The envelope, which we constructed in the example with $\beta = 1/2$ (Diagram 9), thus gets dominated, if better techniques exist in the upper right corner of the wage curve box with probability $\pi^\star > 0$. A baffling conclusion follows: If $\pi^\star$ is given and s is small, a neoclassical envelope similar to the one in Diagram 9 may arise, but if s is increased, the chance that a better technique is adopted rises, and there will be a kind of phase transition to an envelope resembling that of Diagram 3. Numerical experiments that illustrate this effect will be shown in Section 6 (Diagram 12). However, such a phase transition will not take place, if the better techniques are too sparse or if $\delta^\star$ diminishes too fast, as s grows. If, in the proof above, $\delta^\star = \frac{\delta}{s}$ is replaced by a $\delta^{\star\star} = \delta/s^2, \pi^\star \to 0$ for $s \to \infty$. Hence the importance of empirical evidence. The neoclassical case is not impossible, but improbable.

Non-linear wage curves

We have generalised Proposition 2 by going beyond the assumption of a uniform probability distribution; we now use the central assumption to make a step towards the inclusion of non-linear wage curves. The reader may have noticed that the results of Proposition 1 hold also if any two wage curves that appear on the envelope intersect only once, at a switch point on the envelope. This assumption will be fulfilled, if reswitching and reverse capital deepening can be ruled out, following the Appendix in Schefold (2016), that is, if the dimension of (or the number of commodities in) the system is sufficiently large and a boundedness condition holds. If the wage curves that appear on the envelope intersect at most once, we call them disentangled. Since this may be expected, Proposition 1 is general and does not depend on the linearity of the wage curves in an essential way. Nor does it depend on the distribution of the $w^\sigma(0)$ and the R_σ. Proposition 1 only depends on the pairings of the $w^\sigma(0)$ with the R_σ, and if these are not independent, if they are not equally probable, we have $\beta > 1$ or $\beta < 1$. An extension of those cases to non-linear wage curves then is still possible.

Maintaining the central assumption at first and supposing wage curves that have no inflection points and intersect only once on the envelope, we get at once that Proposition 1 holds. Propositions 2 and 3 are in question because of Wicksell effects. Nonetheless, we have not only the expectation of the number of switchpoints, equal to or lower than $\ln(s)$, but we can also make a statement about the distribution of the switch points on the envelope. It is, however, weaker than the one contained in the Theorem of Section 4. As in the proof of Proposition 1, we estimate an upper limit ω for the number of switch points by looking at the envelope 'from above', but we now divide the procedure in two steps, adding up to ω' the probabilities that $w^1, \ldots, w^{\vartheta s}$; $0 < \vartheta < 1$; is on the envelope; ω'' is the sum of the remaining probabilities up to s. Hence

$$\begin{aligned}\omega &= \omega' + \omega'' \cong \ln s,\\ \omega' &= 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{\vartheta s} \cong \ln(\vartheta s) = \ln s + \ln \vartheta,\\ \omega'' &= \omega - \omega' \cong -\ln \vartheta = \ln \frac{1}{\vartheta}.\end{aligned}$$

It follows that a tendentially infinite number of wage curves that appear on the envelope will start from the ordinate at $w^1(0), \ldots, w^{\vartheta s}(0)$, while only a finite number will make it on the envelope of the infinite number of wage curves $w^{\vartheta s+1}, \ldots, w^s$, as s tends to infinity. A symmetric argument can be made by analysing how the wage curves start from the abscissa. By varying ϑ, one finds that the wage curves on the envelope cluster at $r = 0$ and at $r = R_{\sigma_s}$. By contrast, there is only a finite number of wage curves appearing on the envelope with $w^\sigma(0) \in (0, \vartheta s)$ and/or $R_\sigma \in (0, \vartheta R_{\sigma_s})$. It is remarkable that this number depends only on ϑ, not on s. The wage curves appearing on the envelope can therefore cluster only at the ends. Wage curves starting near the ends are likely to appear in the middle part of the envelope, if neoclassical Wicksell effects are sufficiently weak.

If the central hypothesis is replaced by the assumptions $P(R_\sigma > R_{\sigma-1}, \ldots, R_\sigma > R_1) = 1/\sigma^\beta$, $\beta > 0$, $\beta \neq 1$ for all σ, we get, if the wage curves are disentangled and have no inflection points, with the same subdivision for the sequence of wage curves as above:

$$\omega = \omega' + \omega'' = 1 + \frac{1}{2^\beta} + \frac{1}{3^\beta} + \cdots + \frac{1}{s^\beta} \cong \frac{s^{1-\beta} - 1}{1-\beta},$$
$$\omega' = 1 + \frac{1}{2^\beta} + \cdots + \frac{1}{(\vartheta s)^\beta} \cong \frac{(\vartheta s)^{1-\beta} - 1}{1-\beta},$$
$$\omega'' = \frac{1}{(\vartheta s)^\beta + 1} + \cdots + \frac{1}{s^\beta} \cong \frac{s^{1-\beta}}{1-\beta}\left(1 - \vartheta^{1-\beta}\right).$$

In this case, ω'' tends to infinity, and we cannot prove that the clustering takes place exclusively at the ends, as $s \to \infty$. We can use Zambelli's (2017) diagram once more to illustrate the possibilities. It shows the envelope of the techniques derived from the input-output tables of 30 countries with 31 sectors, hence $s = 30^{31} \approx 6.2 \cdot 10^{45}$ and 63 wage curves appear on the envelope, with 62 switch points. If we make the central assumption and apply the formula of the Theorem, we obtain $\frac{2}{3}\ln(30^{31}) \cong 70$, which is plausible. We can also ask what β would have to be, if the formula for ω above is to hold. It implies a β quite close to 1. On the one hand, this justifies the use of the logarithmic approach with its logical stringency. On the other hand, one sees no clear clustering of the wage curves at the ends of the envelopes, which suggests that β is not exactly equal to one.

We combine the analyses of linear and non-linear wage curves by calling the line segment connecting $w^\sigma(0)$ and R_σ the shortcut of the wage curve $w^\sigma(r)$. How close a wage curve is to its shortcut depends on the Wicksell effect which in turn depends on the numéraire. Clearly, the statements of Proposition 5 may be assumed to hold for the anchor points of the shortcuts. Under the conditions of Proposition 5, the shortcuts will tend to the diagonal (Proposition 3), but Wicksell effects may hide the fact.

The following Table 1 summarizes the results in a stylized form, assuming that the s techniques are, looking form above, characterised by $P(R_\sigma > R_{\sigma-1}, ..., R_\sigma > R_1) = 1/\beta, 0 \leq \beta$, and that the wage curves are disentangled. Moreover, we assume, where possible, a uniform probability distribution in the upper right corner of the wage curve box, with $\delta \geq 0$, such that $\delta^\star \geq \delta/s$ (see Proposition 5). The upper bound for the expected number of wage curves is ω, and $\Omega = \omega/s$. The ratio m/s measures the upper bound for the distance of the shortcuts on the envelope from the diagonal of the wage curve box as derived above for the straight wage lines (Propositions 2 and 5).

β	ω	Ω	$\lim\limits_{s\to\infty} \Omega$	$\lim\limits_{s\to\infty} m/s$	$\delta^\star$
$\beta > 1$	$\dfrac{1}{\beta - 1}$	$\dfrac{1}{s(\beta-1)}$	0	undefined	0
$\beta = 1$	$\ln s$	$\dfrac{\ln s}{s}$	0	0	$1/s$
$0 < \beta < 1$	$(1-\beta)^{-1}(s^{1-\beta} - 1)$	$(1-\beta)^{-1}(s^{-\beta} - 1)$	0	0	$\geq \delta/s$
$\beta = 0$	s	1	1	undefined	0

Table 1: Summary of results

The attentive reader of this paper will have understood that good techniques must eventually dominate the envelope, if there is a uniform distribution of the $w^{\sigma}(0)$ and the R_{σ} on the axes and if there is a uniform probability distribution σ_{τ} for the anchor points (case $\beta = 1$). Only a finite number of curves appear on the envelope, if $\beta > 1$. This finite number can be large and is equal to $1/\epsilon$, if $\epsilon > 0$ is small and $\beta = 1 + \epsilon$. Convergence to the diagonal cannot happen in the strict sense of the word. Indeed, there is no room for the uniform distribution, for that would imply an infinity of wage curves on the envelope, contradicting what we found for $\beta > 1$. Ignoring more complex outcomes, we simplify and put $\delta^{\star} = 0$ in Table 1. It is trivial that no dominance arises, if $\beta = 0$, where each diminution of w^{σ} with increasing σ implies an increase of R_{σ} with certainty: the pure neoclassical case. Here there is clearly no room for the uniform distribution of probabilities in the upper right corner of the wage curve box either; we must put $\delta^{\star} = 0$. The interesting question concerns the transition with $0 < \beta < 1$. If $\delta = 0$ or $\delta^{\star} = \delta/s^2$, the case is neoclassical. But the probabilities in the upper right corner may exceed δ/s (as long as the probability sums on rows and columns add up to one, of course). As long as s is not large, the envelope may look quite neoclassical, but the relevant separation between the domains is not between $\beta = 1$ and $\beta < 1$, although there is a qualitative difference also there. If $\beta = 1$ and $s \to \infty$, the wage curves cluster at the end points near $r = 0$ and $r = R_{\sigma_s}$ and the number of wage curves in between is finite, if they are linear (Kersting‘s theorem), while such clustering need not take place, if $\beta < 1$, as we have seen. But, from the economic point of view, it seems more important that Ω, the density of the wage curves, tends to zero with $s \to \infty$ as soon as $\beta > 0$ and, more important still, good techniques must come to dominate at very high levels of s and no room is left for the substitution of capital and labour, hence for the marginal productivity theory of distribution, unless – somewhat ironically – Wicksell effects destroy the simplicity of the conclusion. So the relevant divide appears to be between $\beta = 0$ and $\beta > 0$, if $\delta^{\star} \geq \delta/s > 0$.

6 Numerical experiments and empirical results

Only a few results of the investigation can here be presented. The empirical data were collected and evaluated by an able student, Jakob Kalb, whom we should like to thank. We begin with the numerical experiments. The techniques are represented by randomly generated coordinate variables for the anchor points in a box, which is of given side length for the cases with a uniform distribution on the axes, while expectation and variance are given for the cases with normal distribution.[22] The diagrams show those corresponding linear wage curves which appear on the envelope; inferior curves are suppressed. The number of randomly generated techniques can here be interpreted as the size of the sample or as the number of iterations of the programme. The experiments vary on the one hand according to assumptions about the distribution on the coordinates: It may be uniform or normal. On the other hand, an assumption is made about the correlation between the maximum wage rates (the values on the ordinate) and the maximum rates of profit (the values on the abscissa). The correlation replaces the assumption about β in Section 5. Neoclassicals seem to expect a negative correlation. Our empirical data suggest the opposite, as will be explained later. A negative correlation means that the probability for the anchor points to be near the diagonal is higher than that to be near the lower left or upper right corner of the wage curve box. If all anchor points are on the main diagonal, an envelope of hyperbolic shape is implied.

We present four diagrams according to this scheme, juxtaposing the two kinds of distribution with a zero or an extremely negative correlation in order to investigate where the theses presented in this paper may go wrong. The following Table 2 shows the four cases considered:

Distribution	Correlation	
uniform	0 (case A)	-0.99 (case C)
normal	0 (case B)	-0.99 (case D)

Table 2: Overview of the four cases considered.

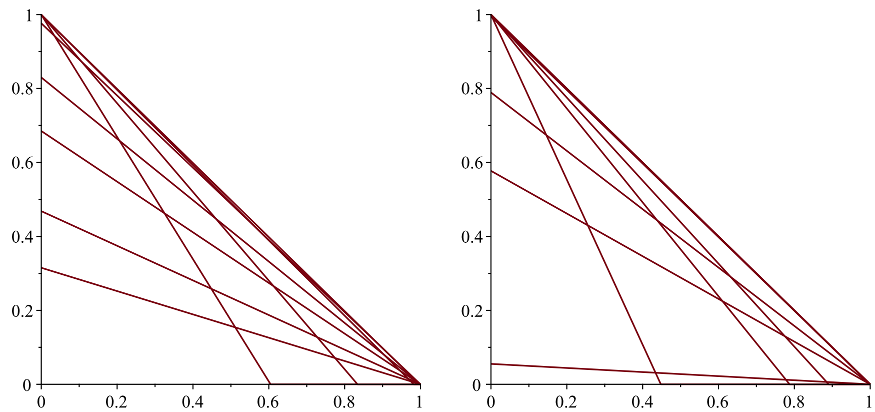


Diagram 10 (case A): 10^5 and 10^8 iterations. With a uniform distribution and no correlation, the number of wage curves is from the start small in the middle range. After 10^8 iterations, the diagonal dominates and the wage curves cluster at the ends (compare Diagram 3 above).

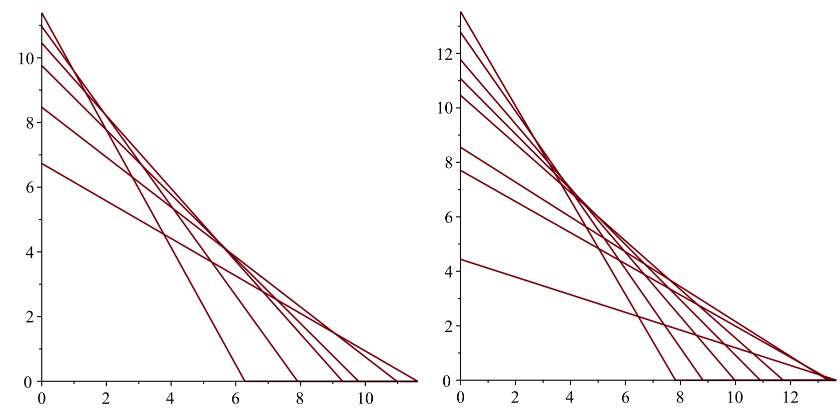


Diagram 11 (case B): 10^5 and 10^8 iterations. The change to a normal distribution without a correlation means that there are fewer anchor points in the upper right corner of the wage curve box, so that diagonal dominance is less pronounced, but also fewer appear along its upper and right side (EF and FB in Diagram 2 above). Hence there is no clustering near the ends, the envelope gets nearly straight and consists of a modest number of curves even with 10^8 iterations. It is the case closest to the empirical data below, except that we observed a positive correlation, which leads to an envelope, which is more straight (compare Diagram 22).

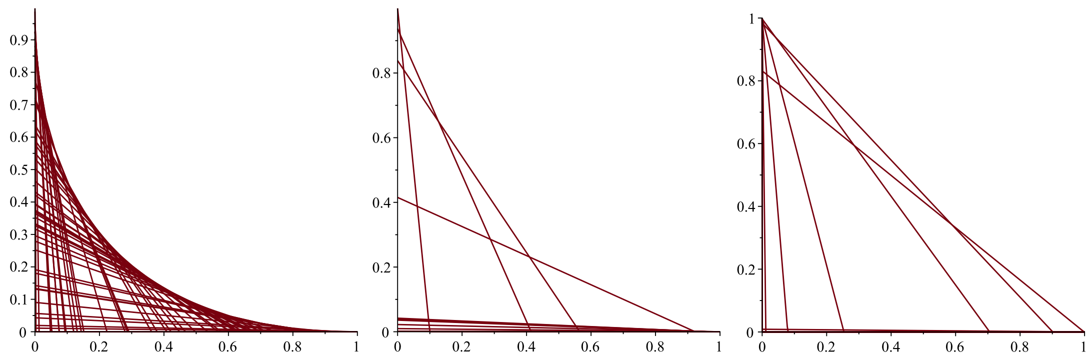


Diagram 12 (case C): 50, 500 and 10^5 iterations. Case C exhibits a striking phase transition. With 50 iterations, one gets a nearly neoclassical envelope, thanks to the strong negative correlation, but it breaks down and a large number of wage curves clusters at the ends (visible as thicker coordinate axes), while only a few wage curves appear in the middle, after only 500 iterations A second transition occurs after 10^5 iterations: the envelope approaches the diagonal and consists – except at the ends – of only two wage curves.

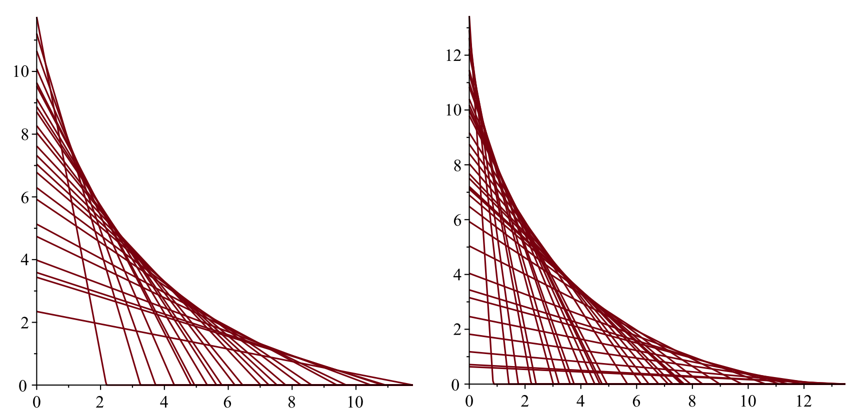


Diagram 13 (case D): 10^5 and 10^8 iterations. The neoclassical case can arise, if the distribution is far from uniform and the correlation is strongly negative. A scatter plot diagram for the anchor points shows them clustered near the main diagonal.

Before looking at the empirical data, we have to confront the fact that the efficiency frontier – the envelope of the wage curves – seems not to get reached in practice. The input-output tables of, as above, 10 countries with 100 sectors yield 10^{100} techniques, which all belong to a potential, of which only 10 are actually realised by the countries concerned. According to both neoclassical and classical abstract theory, production takes place at the efficiency frontier, but no country is actually found there (Zambelli et al., 2017), for the following main reasons: Input coefficients are based on monetary measures of physical input requirements at a level of aggregation which is high relative to the concrete technical methods used; to group techniques, one would need a fuzzy logic. Techniques e.g. for car production, which look similar, but are different in input-output tables in France and Germany, could then be classed as the same. With such a grouping of techniques, sufficiently broadly defined, the number of techniques could be reduced drastically and one might find actual techniques on the envelope after all. Then there are problems of information. If better techniques are to be imitated, the techniques may be tied to institutions that change at best slowly or to geographical conditions that are definitely not transferable (we noted such specificity of techniques in Section 3). Transitions between techniques take time and may need several steps (steel must be produced, before tankers can be built). Scale effects and joint production may also play a role (with joint production, the dynamic of adoption of better techniques because of surplus profits is different from that which holds with single production). Nonetheless, for lack of alternatives, we continue to take combinations of methods, represented as sectors of input-output tables, as our inventory of techniques. Note that if we were to restrict our attention to 'transferable techniques' – means of production such as lorries and computers are obviously transferable – and if only one in one hundred techniques are transferable, there remain in the example still an enormous number: 10^{98}.

If we place the potential techniques of e.g. ten countries as anchor points in a wage curve box, special attention will be paid to the ten actual techniques and to what we shall call the graspable potential, defined as the anchor point given by the maximum output per head and the largest maximum rate of profit of the ten actual techniques. The graspable potential turns out to be a rough indicator of the upper right corner of the wage curve box and therefore of an area, where anchor points are found that represent wage curves on or near the envelope.

We turn to the empirical investigation proper. The input-output matrices result from the input-output tables of several countries for the year 2014, which were published in the release "World Input-Output Database (WIOD)" of the year 2016. On these data see Timmer et al. (2015) and Timmer et al. (2016). Two sectors ("activities of households as employers" and "activities of extraterritorial organizations and bodies") were eliminated as not significant; the remaining 54 sectors were used for the calculation of the anchor points. As usual, the input-output coefficients result from the division of the monetary flows of intermediate goods by the monetary value of the gross output of each sector. We adopt the fiction that each country is self-sufficient by adding the coefficients for imported goods in each sector to the corresponding domestically produced inputs; exports then are part of the surplus[23]. The calculation of the labour vectors is based on the "Socio Economic Accounts" of the WIOD database; they were normalised in the same way as the input-output coefficients. The common net output of all countries is obtained by forming the average of the countries' sectoral "value added"-entries in their input-output tables. This net output is chosen as the numéraire and kept constant in order to allow for a meaningful interpretation of the countries' respective maximum wage rates as output per head or, more precisely, per labour hour. The calculations of output per head and of the maximum rate of profit for each country, derived from the Frobenius-eigenvalue of the matrix, then follow the exposition of Notes 1, 2, 3 and 11 of the paper.

The following Diagram 14 shows the calculation of anchor points in a scatter plot for the input-output tables of France and Germany. The correlation coefficient $\rho = -0.086$ is small, but statistically significant ($p \approx 6.8 \times 10^{-18}$). Kendall's rank correlation was also calculated, $\tau = -0.0576$. This coefficient is statistically significant, too ($p \approx 5.4 \times 10^{-18}$).

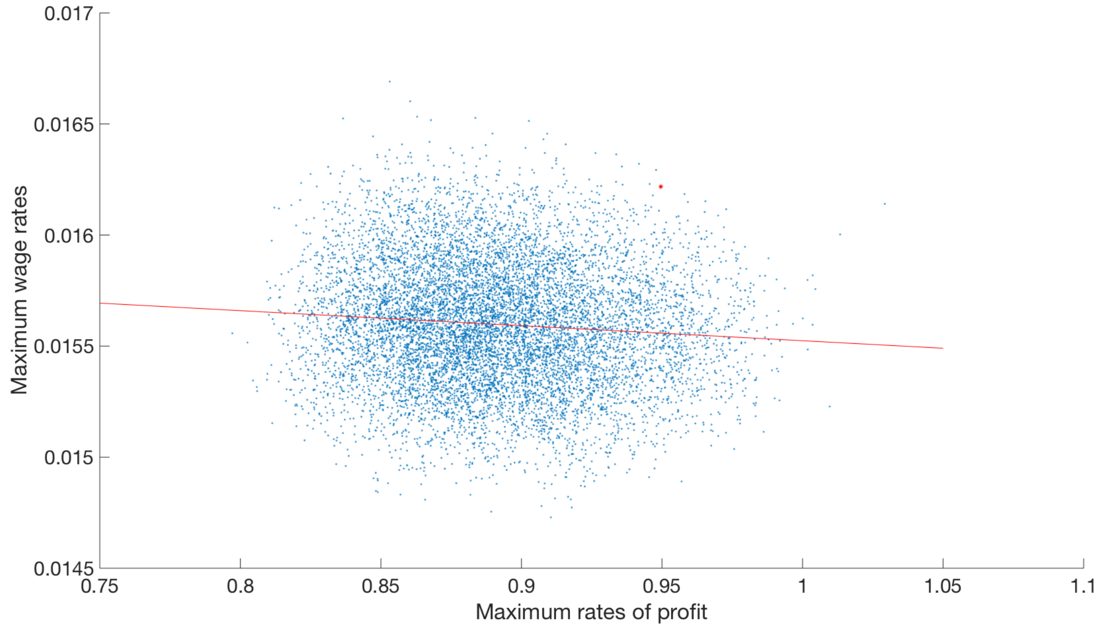


Diagram 14: 10000 anchor points calculated on the basis of French and German input-output tables. Each blue dot corresponds to an anchor point; the red line is the result of a simple linear regression. The red star marks the graspable potential, in this case consisting of the French maximum wage rate and the German maximum rate of profit.

We add the histograms of the maximum wage rates and the maximum rates of profit in Diagram 15.

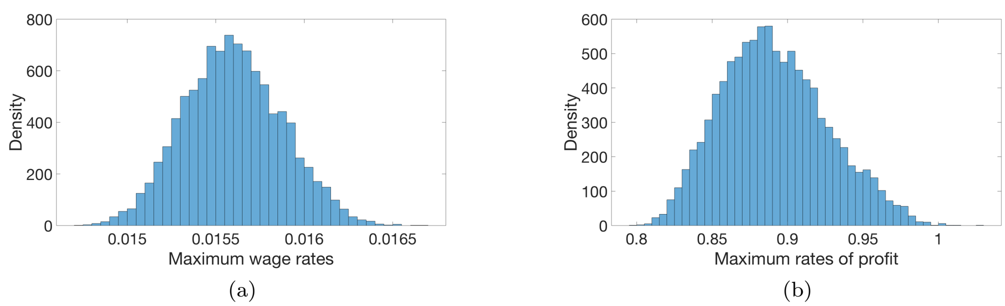


Diagram 15: Histograms of the maximum wage rates (a) and the maximum rates of profit (b), based on French and German input-output tables.

The graspable potential is within the cloud of the scatter points and may indicate techniques that dominate the envelope of the shortcuts corresponding to the anchor points. The marginal distributions of the maximum wage rates and maximum rates of profit, however, are clearly not uniform, but closer to normal distributions. This was confirmed by means of Q-Q plots for the distribution of the wage rates and the rates of profit. The plot shows how the standard normal distribution correlates with the empirical distribution, each measured by means of quantiles. A perfectly normal empirical distribution therefore would result in a linear relationship. The diagrams 16 and 17 confirm that we get a good fit.

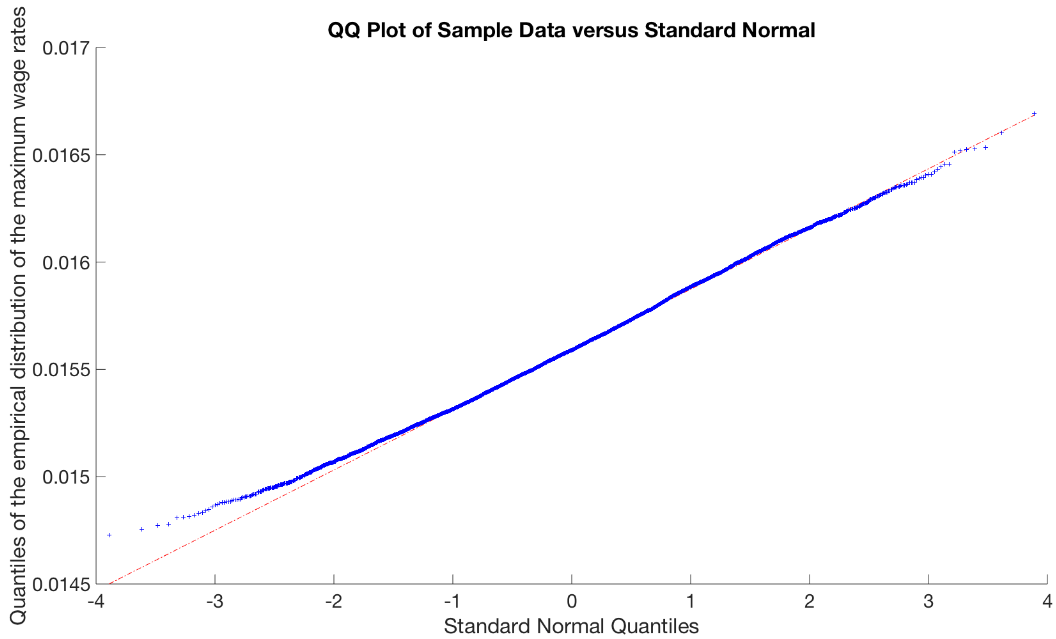


Diagram 16: Q-Q plot for the maximum wage rates.

The scatter diagrams are based on a random selection of 10000 techniques or combinations of French and German methods, of which we have 2^{54}. The analysis of the Q-Q plots then was complemented by a Shapiro-Wilk test (Shapiro and Wilk, 1965) on the basis of 5000 wage and profit rates taken from the 10000 because of a limit in the program. A p-value of $p \approx 5.5 \times 10^{-7}$ was obtained for the maximum wage rates and of $p \approx 2.2 \times 10^{-16}$ for the maximum rates of profit. The small p-values demonstrate that the variables do not follow a normal distribution strictly, but the sample is large (5000); the Shapiro-Wilk test may be excessively restrictive and the Q-Q plot more telling for our purposes.

In the next step, three countries, the United States, Great Britain and Italy, were added. The linear regression in Diagram 18 for the anchor points now reveals a slightly positive correlation between $w(0)$ and R.

The correlation coefficient is $\rho = 0.1989$, the rank correlation is $\tau = 0.1290$. The histograms and the Q-Q plots look similar to what we found in the two country case. Two observations can be made: There seem to be fewer techniques with very small maximum wage rates than would correspond to a normal distribution, while we have a deviation from the normal distribution in the case of the maximum rates of profit at the upper end: There seem to be more techniques with a very high maximum rate of profit in each of the five countries than a normal distribution would imply. The graspable potential for each country was calculated. A table was made for an anchor point that seems to correspond to a technique with a wage curve on the envelope, according to visual inspection; it shows the best method used in each sector and the country, in which that

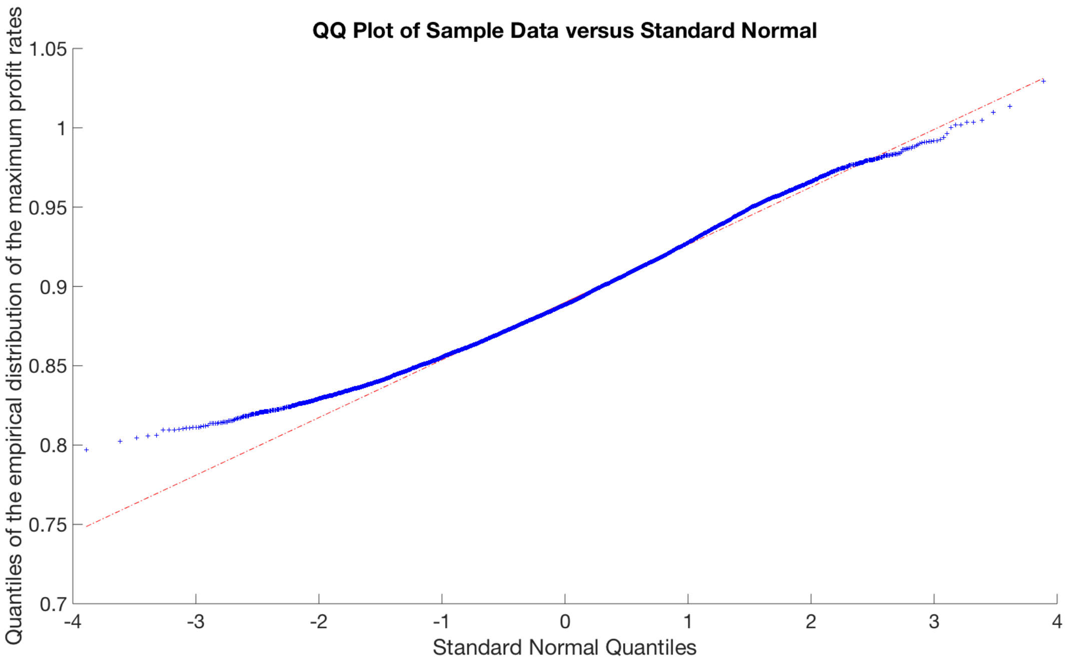


Diagram 17: Q-Q plot for the maximum rates of profit.

method is in actual use. All the five countries contribute methods to this technique in a colourful mixture, which is not easy to interpret. It is a potentially dominating technique, but it is sure to be on the envelope only in the given sample of 10000 techniques and the calculation does not refer to the envelope of the wage curves, but to their shortcuts. There is no room to represent it here.

These results have proven to be robust to variations in the number of considered countries and years. To check the robustness of our results, we replicated our calculations for more countries and different years, using a pre-processed data set published by Zambelli and referred to in Zambelli (2018). This data set allowed us, after appropriately adjusting the national input-output matrices and labour vectors to a suitable definition of net output and our notation, to calculate anchor points of techniques generated by randomly combining methods of production of ten countries (Canada, Finland, Spain, Netherlands, Austria and the five countries from above) for each year from 2001 until 2011. In all cases, the correlation coefficient between $w(0)$ and R is positive and exceeds 0.3. The negative correlation of the two country case therefore clearly presents itself as an exception from the rule. Furthermore, we observed a negative relation between the number of shortcuts that appear on the envelope and the correlation between $w(0)$ and R. This result was obtained from regressing the number of shortcuts on the envelope on the natural logarithm of the number of techniques, $\ln(s)$, and the correlation between $w(0)$ and R. The regression coefficient for the relation between the correlation and the number of shortcuts on the envelope was negative and statistically significant at the 95% level: A higher correlation between $w(0)$ and R was associated with fewer shortcuts on the envelope in our sample.

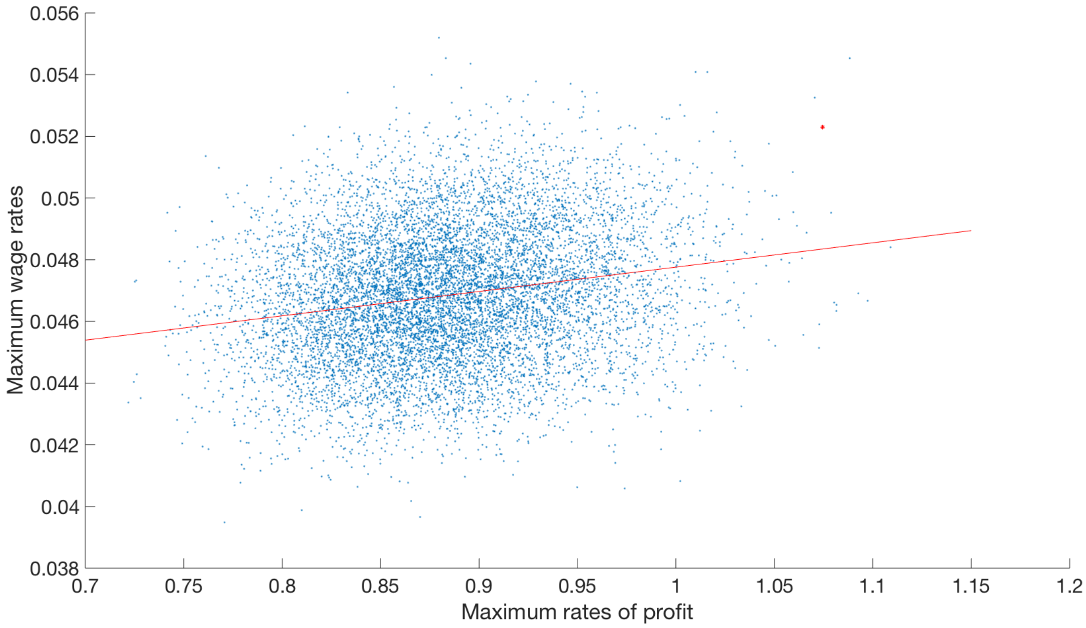


Diagram 18: 10000 anchor points calculated on the basis of French, German, US, British and Italian input-output tables. Each blue dot corresponds to an anchor point; the red line is the result of a simple linear regression. The red star marks the graspable potential in the sense of Section 5, in this case consisting of the Italian maximum wage rate and the US maximum rate of profit. Note that, if the number of techniques is increased (the diagram is based on 10^5 out of 5^{54}), the number of anchor points around the graspable potential will increase significantly, while the graspable potential itself stays in place.

Concluding observations

Because of the complexity of the argument, the paper cannot be concluded without a summary of its main contents. We started from the Cambridge debate, which was concerned with the logical consequences of the fact that wage curves are in general not straight; we recalled declination, non-neoclassical Wicksell effects, reswitching and reverse capital deepening as the main objections raised in the debate against the marginal productivity theory of distribution and the aggregate production function. The critique turned out to be entirely successful, if the production function is understood as a mathematically rigorous theory, but is less pertinent, if it is understood as an approximation to a reality, for which there is no known comprehensive theory. Wicksell effects are numéraire-dependent and wage curves intersect only rarely more than once, if one of the switch points is on the envelope, as has been shown theoretically and empirically; wage curves are disentangled as a rule, and the macroeconomic consequences, if they are not, are small, unless reswitching and reverse capital deepening are frequent. It is a more general problem of this critique that it leads to agnosticism regarding the relationship between distribution and the capital-labour ratio. Joan Robinson, who had initiated the debate, therefore doubted its relevance, and the Cambridge economists pursued a different line in their positive contributions to modern economic theory. They emphasized technical progress and described growth paths according to stylized facts which included constant shares in distribution and a constant capital-output ratio.

Samuelson, in his defence of marginal productivity theory, had assumed that the wage curves of individual techniques were linear, and we have adopted this counterfactual assumption in the first Sections of this paper in order to question Samuelson‘s implicit hypothesis that the productivities of labour and of capital, output per head and the maximum rate of profit for each technique, are in essence ordered inversely.

It was simple to show that, if this hypothesis is not adopted, if instead, all permutations of maximum rates of profit are equally probable, as one descends the envelope of wage curves, the number of wage curves on the envelope rises only slowly, with the natural logarithm as an upper bound so that the number of possibilities of substitution is surprisingly small – too small to verify the neoclassical propositions, as we argued, using plausible numerical examples.

We then translated the assumption made regarding the equiprobability of the permutations into the narrower assumption of a uniform probability distribution and we assumed that there are given bounds for output per head and the maximum rates of profit. If the number of techniques increases under such circumstances, with growing density in the grid representing the anchor points (Diagram 2), the envelope must approach the diagonal of the wage curve box. At that capital-labour ratio, distribution is almost totally indefinite. A more exact calculation shows that the number of wage curves on the envelope then approximates $2/3 \ln s$ and the wage curves will cluster at the ends of the envelopes; only a a few are there in the middle. Although there may be a few switch points, the capital-labour ratio turns out to be essentially determined.

We then widened the assumptions in two different ways:

The simpler and less rigorous approach started again from the possible permutations of maximum rates of profit adopted at a switch point, as one descends the envelope. We supposed that the probability of encountering a switch point, at which the maximum rate of profit rises in such a way that the new wage curve dominates after the switch, equals $1/\sigma^{\beta}$. If such a characterization is possible, it turns out that, if $\beta > 1$, the number of wage curves on the envelope will be finite. The case considered extensively in Sections 3 and 4 corresponds to $\beta = 1$ and the extreme neoclassical case, in which each wage curve becomes part of the envelope, is characterized by $\beta = 0$. So we could portray the transition form the classical perspective, where the capital-labour ratio is essentially predetermined, to the neoclassical case, where it changes smoothly with distribution, as a continuous transition. If the assumption was added that output per head and maximum rates of profit are bounded and that the probability distribution is uniform near the limits, it again turned out that the envelopes would approach the diagonal, for any $\beta > 0$. This was confirmed in numerical experiments, in Diagram 10 (uniform distribution, no correlation) and Diagram 12 (uniform distribution, negative correlation). The case of a uniform distribution with a positive correlation is not shown; it obviously leads to an envelope close to the diagonal.

However, in numerical experiments and in empirical investigations, we also tested another assumption. If the distributions of output per head and of maximum rates of profit were normal, the outlook changed somewhat and the outcome depended strongly on what now replaced the assumption of a uniform probability distribution, that is, the correlation between maximum wage rates and maximum rates of profit.

Three characteristic cases can be distinguished according to numerical experiments, which we now represent by means of the anchor points in Diagrams 19 a, b, c.

In case a, the anchor points are strongly negatively correlated so that they form a cloud of roughly elliptical shape along the main diagonal of the wage curve box, for which, for a given number of techniques, we have a given size. However, the size of the box increases slowly, but indefinitely, with the number of techniques. For a given size, the elliptical cloud must result

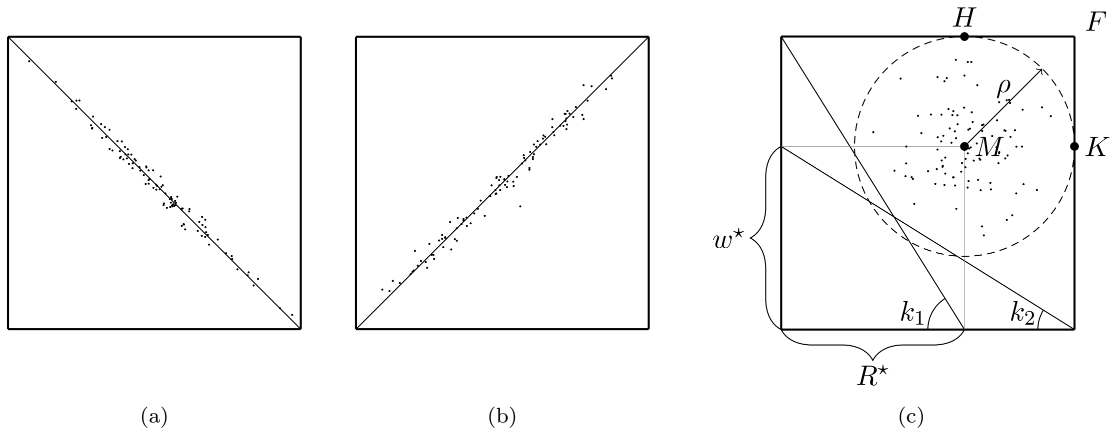


Diagram 19: Three characteristic clusterings of anchor points with a strongly negative (case a), strongly positive (case b) and zero correlation (case c).

in a roughly hyperbolical shape of the envelope of the wage curves: the neoclassical case; it corresponds to Diagram 13.

If there is a strong positive correlation, the anchor points will cluster around the minor diagonal for a given large number of techniques and a given size of the wage curve box. It is intuitively clear that, with increasing size, the envelope of the wage curves will approach the diagonal and may consist of one wage curve only: the best technique.

In between cases (a) and (b), there is case (c) with a correlation equal to zero. Let us stylize the assumption by postulating that the anchor points for a given size fill a circle in the upper right corner of the wage curve box, densely towards the center of the circle, sparsely near its circumference. We assume in accordance with our empirical data that this circle is in the upper right corner of the wage curve box because very inefficient techniques are not represented (all countries considered are similar). The anchor point of the technique with the highest output per head is denoted by H and the anchor point of the technique with the highest maximum rate of profit is denoted by K and F is, as in Diagram 2, the upper right corner of the wage curve box. The center of the circle is M, its radius ρ, and M is at a distance $R^\star$ from the origin along the abscissa and at a distance $w^\star$ from the origin on the ordinate. It is the point of the exercise that this stylized picture corresponds very well to the empirical picture of Diagram 18 and to 25 other similar tests with similar results – with only the difference that has been mentioned: the empirical correlation tends to be positive, favouring the classical result.

Sparse points on the circumference between H and K will be representative of the envelope of the shortcuts of the wage curves. This envelope will correspond to the case represented in Diagram 11. The envelope will remain at some distance from the diagonal, because the anchor points do not approach F. So this case could be said to be in between the pure classical and the pure neoclassical case. Increasing the number of techniques may mean that H, K and F shift outward so that we cannot expect a phase transition to the classical case, as long as no limits are imposed.

Actually, the envelope will be constituted by taking into account also some anchor points slightly inside the circumference, but we may neglect this at the present level of abstraction.[24] Although we get an envelope with substitutions, it will not satisfy the ideal criteria of neoclassical theory. The capital-labour ratios will not vary between zero and infinity, as with an ideal production functioning fulfilling the Inada-conditions, but they can vary only between k_1 and k_2 (see Diagram 19 c): The capital-labour ratios pertaining to the anchor points H and K. The capital-labour ratios, measured along the shortcuts, must lie between $k_1 = \dfrac{w^\star + \rho}{R^\star}$ and $k_2 = \dfrac{w^\star}{R^\star + \rho}$ so that the growth of the capital-labour ratio, as one reduces the rate of profit and increases the wage rate from one extreme to the other, is given and limited by

$$\frac{k_1}{k_2} = \frac{(w^\star + \rho)(R^\star + \rho)}{w^\star R^\star}.$$

It seems to be the case that ρ is not large relative to $w^\star$ and $R^\star$, and if ρ goes to zero, relatively, we get $k_1 = k_2$; the circle is so small as to approach F. This is the case, which, according to the earlier investigation, corresponds to a uniform probability distribution.

Empirically, we found $k_1/k_2 = 1.4377$ for the five country case (for $s = 10,000$). If one smoothes the envelope, the elasticity of substitution, measured locally, must be quite small. A zero elasticity means no substitution and there is one best technique. So the economic result is essentially the same as the one obtained in Sections 3 and 4, where the straightforward assumption of a uniform probability distribution allowed to deduce unequivocal mathematical statements. According to our empirical findings, the correlation between the distributions of output per head and maximum profit rates is not zero but positive; one gets an intermediate case between (b) and (c).

An elasticity of substitution of around one seems to be favoured in most neoclassical studies and it must be large enough for important applications such as Böhm-Bawerk's idea that an elevation of the real wage above its natural level could be countered by an increase of the capital-labour ratio, using known techniques. We could conclude with the statement that the elasticity of substitution is too small for the essential applications of neoclassical theory.

But this is not the end of the story, for the idea that the techniques are normally distributed without bounds may be questioned. There is an irony in using the normal distribution, which descends from Gauss's theory of errors with the focus on the expectation. But optimisation implies an analysis of extreme values, which would be misses, if the errors represent deviations from a target. If we have, as in our example, ten countries and 10^{54} techniques, only ten techniques, those of the countries concerned, are actual. Among the other techniques, many may be specific and not transferable – a problem which we discussed – some may not be realisable, at the activity levels derived from the given output-level, because of natural constraints. Unfortunately for theorists of whatever persuasion, the techniques that appear best on paper by being close to or on the envelope are the most suspicious by being farthest away from those actually used. So there may be limits below H and to the left of K in Diagram 19c.

A limit of some plausibility is given by the graspable potential. It pairs the productivity of labour reached by the country most successful in this regard with the corresponding productivity of capital reached by, in general, another country, and chances are that techniques remain to be realised, by combining existing methods, so that one technique reaches both these limits.

This means to assume that the box is restricted in the upper right corner by a point that may be identified with the graspable potential. This point coincides with the upper right corner F in Diagrams 2 and 19c. We assume that the corresponding triangle DFG of Diagram 2 is not empty. Since we are dealing with normal distributions, we restrict the box in the lower left corner by assuming that point A (Diagram 2) now corresponds to the expected values of $w^\sigma(0)$ and R_σ (anchor points to the left and below are dominated). Whatever the correlation c, $-1 < c \leq 1$, the density then diminishes, as one moves up and to the right from A to F. We assume that the probability density for anchor points does not drop to zero in approaching F for given s and that it does not diminish as s increases. Let the minimum density be smaller than the maximum density by a (possibly small) factor $\xi > 0$.

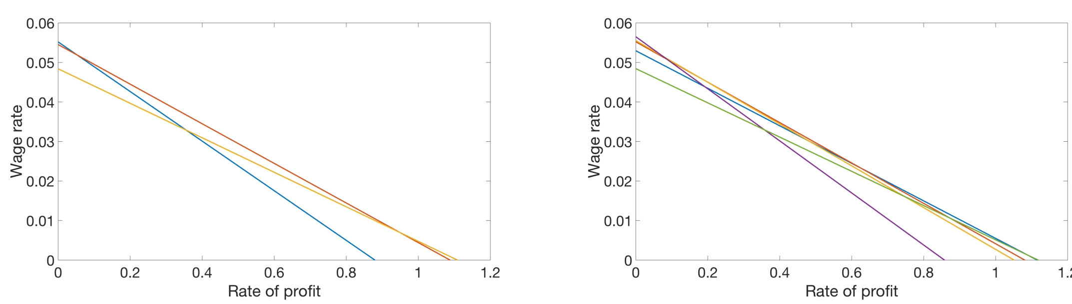
Proposition 6
Under the stated assumptions, there is a probability $\pi^{\star\star}$ *of at least* $1 - 1/\sqrt{e^{\gamma\xi}}$ *that the envelope moves to the diagonal with* $s \to \infty$, γ *given* $(0 < \gamma < s)$.

Proof. The probability π_1 that we find an anchor point in DGF, which also is in $ABEF$, is at least equal to the ratio of the minimal density times the area of DFG, divided by the maximum density times the area of $ABEF$. Hence $\pi_1 \geq \frac{\xi m^2}{2}/s^2$. We assume, as in the strong case, that m grows with s, but more slowly, according to $m = \sqrt{\gamma s}$. The probability π_2 that all s anchor points are not in DFG, but in $ABEF$, is $\pi_2(s) = (1 - \pi_1)^s \leq (1 - \frac{\xi}{2}\frac{\gamma}{s})^s$. Then the probability $\pi^{\star\star}$ is $\pi^{\star\star} = \lim_{s\to\infty} 1 - \pi_2(s) \geq \lim_{s\to\infty} (1 - \frac{1}{s}\frac{\xi\gamma}{2})^s = 1 - \frac{1}{\sqrt{e^{\gamma\xi}}}$. We again arrive at the results found for the strong case, but now by Euler's definition of e, not by the logarithm.

The existence of bounds is plausible not only in the case of uniform, but also of normal distributions, although they are difficult to identify. Fortunately, the empirical results give a clear picture also if no bounds are assumed or imposed. The envelopes for all the samples of anchor points collected in this paper have been examined – they are, of course, envelopes of the shortcuts corresponding to the anchor points. A program first eliminates all anchor points that are dominated by some other anchor points; the number of those not dominated is $\widetilde{Z}$. $\widehat{Z}$ denotes the number of shortcuts that actually appear on the envelope. Of course, $\widehat{Z} \leq \widetilde{Z}$, since anchor points and shortcuts may be dominated, even if they are not dominated by an individual anchor point or shortcut, but by a combination. The result is shown in Table 3.

As one can see, $\widehat{Z}$ is below $(2/3)\ln s < \ln s$ in all cases. The assertions made about the number of wage curves on the envelope on the basis of the strong case therefore are not only confirmed: it seems that the formula $(2/3)\ln s$ derived by Kersting for the (on the margin) uniform distribution must be replaced by a lower estimate (the theoretical extension of Kersting's theorem to the case of normal distributions, which seem a good approximation for reality, has not yet been found). The positive correlation, here found empirically, might play a role.

The envelopes of the shortcuts themselves turn out to be bundled. By this we mean that there is little variation of the capital-labour ratios (see as examples, Diagrams 20 a and b) and that was, ultimately, the simple point to be made.



(a) (b)

Diagram 20: Envelopes of (a) 10,000 shortcuts and (b) 100,000 shortcuts. Both plots are based on input-output data of five countries (GER, FRA, USA, GBR and ITA) for the year 2014, corresponding to the (a) second and (b) third entry of Table 3.

No. of countries	Year	s	$\widetilde{Z}$	$\widehat{Z}$	Rank corr.	Correlation	$\ln(s)$	$2/3ln(s)$
2	2014	10000	14	3	-0.06	-0,09	9.21	6.14
5	2014	10000	4	3	0.13	0,.0	9.21	6.14
5	2014	100000	7	5	0.13	0.20	11.51	7.68
5	2011	10000	6	4	0.23	0.35	9.21	6.14
10	2011	10000	6	4	0.21	0.33	9.21	6.14
10	2011	100000	8	3	0.21	0.32	11.51	7.68
10	2010	10000	5	3	0.22	0.35	9.21	6.14
10	2010	100000	4	3	0.23	0.35	11.51	7.68
10	2009	20000	4	3	0.26	0.39	9.90	6.60
10	2009	80000	8	3	0.26	0.39	11.29	7.53
10	2008	15000	4	3	0.26	0.39	9.62	6.41
10	2008	75000	8	4	0.25	0.39	11.23	7.48
10	2007	10000	8	4	0.27	0.42	9.21	6.14
10	2007	60000	10	5	0.28	0.43	11.00	7.33
10	2006	12000	5	3	0.26	0.40	9.39	6.26
10	2006	50000	5	3	0.28	0.42	10.82	7.21
10	2005	15000	3	3	0.27	0.42	9.62	6.41
10	2005	55000	9	5	0.27	0.42	10.92	7.28
10	2004	10000	10	5	0.25	0.39	9.21	6.14
10	2004	50000	6	4	0.26	0.40	10.82	7.21
10	2003	10000	3	3	0.27	0.42	9.21	6.14
10	2003	50000	5	4	0.27	0.41	10.82	7.21
10	2002	10000	7	4	0.28	0.43	9.21	6.14
10	2002	45000	7	5	0.28	0.43	10.71	7.14
10	2001	10000	2	2	0.27	0.41	9.21	6.14
10	2001	60000	8	4	0.27	0.41	11.00	7.33

Table 3: The numbers of shortcuts on the envelopes, the correlations and the predictions for the numbers by means of the formulas $\ln s$ and $(2/3)\ln s$. Data from WIOD and pre-processed data from Zambelli (2018); see Section 5 for more details on the data.

We confirm the empirical results by coming back to the numerical experiments. We considered the possibility of cutting off extreme outliers of anchor points by restricting the analysis to the graspable potential. This has no direct equivalent in the numerical experiments, but if extreme values are cut off, the only case which had resulted in a clear tendency to a smooth envelope of hyperbolical shape, on the basis of a normal distribution, coupled with a strong negative correlation (Diagram 13), straightens considerably (Diagram 21).

We have found good evidence for a normal distribution, but with a positive, not a negative, let alone an extremely negative correlation. Empirically, the correlation is near 0.4. The corresponding numerical experiment (Diagram 22) shows how in this case the number of lines on the envelope is small after 10^5 iterations and diminishes, instead of rising, as the number of techniques in the sample increases to 10^8. The envelope consists of only three line segments and is strikingly similar to the empirical result of Diagram 20.

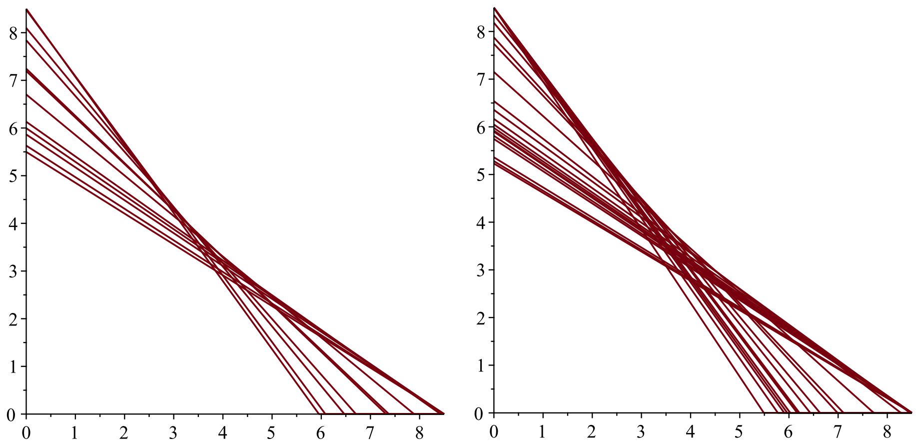


Diagram 21: Numerical experiment. 10^5 and 10^8 iterations. The correlation is -0.99, the distribution is normal, all as in Diagram 13, but upper limits have been imposed for both coordinates.

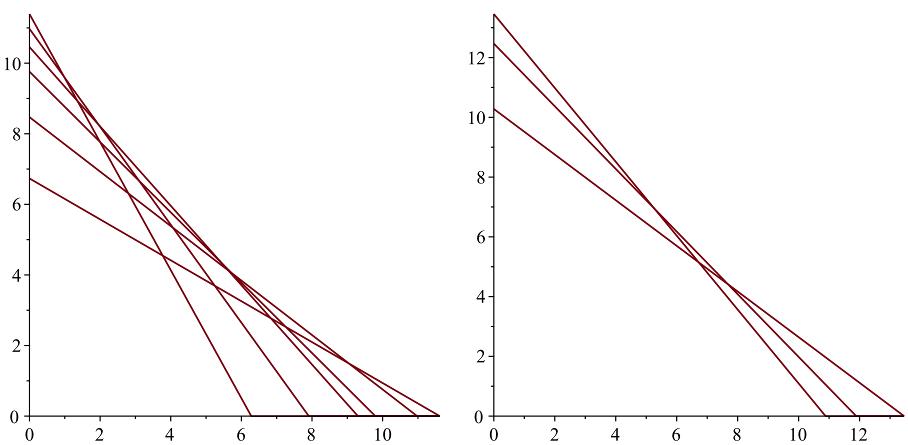


Diagram 22: Numerical experiment. Normal distribution. Correlation 0.4. 10^5 and 10^8 iterations.

Notes

[1]A technique is here usually given by a semi-positive, productive, indecomposable input-output matrix $\mathbf{A}$ for the production of n goods, $\mathbf{A} = (a_{ij})$, where a_{ij} denotes the amount of commodity j used for the production of a unit of commodity i (industries on the rows, following Sraffa's notation). Each industry i uses labour l_i; the labour vector is $\mathbf{l}$. There is a uniform rate of profit or interest r, a uniform wage rate w; the economy is in a stationary state, the wage is paid *ex post*, and prices are given by

$$\mathbf{p} = (1+r)\mathbf{A}\mathbf{p} + w\mathbf{l} = w(\mathbf{I} - (1+r)\mathbf{A})^{-1}\mathbf{l}.$$

[2]There is a vector $\mathbf{d}$; it represents the basket of goods, of which the net product is composed. This vector is also used as the numéraire, hence $\mathbf{dp} = 1$. Prices and the wage rate are then determined as functions of r, $\mathbf{p}(r)$ and $w(r)$. The wage curve $w(r)$ is monotonically falling between $r = 0$ and a maximum rate of profit R, where $w = 0$. If $\hat{\mathbf{p}} = \mathbf{p}/w$ are prices in terms of the wage rate, one has $1 = \mathbf{dp} = \mathbf{d}\hat{\mathbf{p}}w$, hence $w = 1/\mathbf{d}\hat{\mathbf{p}}$.

[3]We denote output per head by $y = \mathbf{dp}/\mathbf{ql}$, $\mathbf{q}$ being the activity levels to produce $\mathbf{d}$:

$$\mathbf{q}(\mathbf{I} - \mathbf{A}) = \mathbf{d}.$$

Since all income goes to wages at $r = 0$, we have $y(0) = w(0)$. Since output is the numéraire, $\mathbf{dp} = 1$ and $y(r) = w(0) = \mathbf{dp}/\mathbf{ql} = 1/\mathbf{ql}$ is independent of r, and $y = rk + w$, k intensity of capital, $k = K/L$; $K = \mathbf{qAp}$ total capital, $L = \mathbf{ql}$ labour employed. Hence $k = (1/r)(y - w)$. Clearly, k is given by $\tan\varphi$ at each rate of profit in Diagram Note 3, showing the wage curve of a given technique:

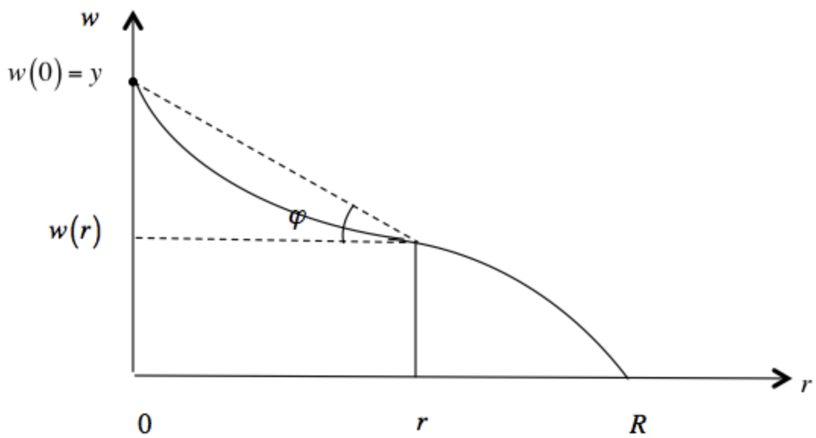


Diagram Note 3: Wage curve of a given technique with capital-intensity $\tan\varphi$ at r.

[4]If there are several techniques producing the same goods, each will be characterized by its own wage curve. The profit-maximizing technique will at each r be that of the wage curve on the envelope. Wage curves intersecting on the envelope have generically all methods in all industries in common, except in one: the method is changed only in one industry. The intersection where the method change occurs is called a switch-point. If the wage curves happen to be straight lines as in the case of the wage curves w^3 and w^4 in Diagram Note 4, the capital-intensity does not change along the wage curve, and increasing r means the adoption of wage curves of lower capital intensity, as r rises and w falls at the switch-point D on the envelope. This inverse relationship between the capital-intensity and the rate of profit does not hold between A and B along wage curve w^2 because of its concave curvature: the intensity of capital rises between A and B. This is a non-neoclassical Wicksell-effect. Moreover, the intensity of capital rises at B in the transition from w^2 back to w^1 on the envelope (so-called reswitching).

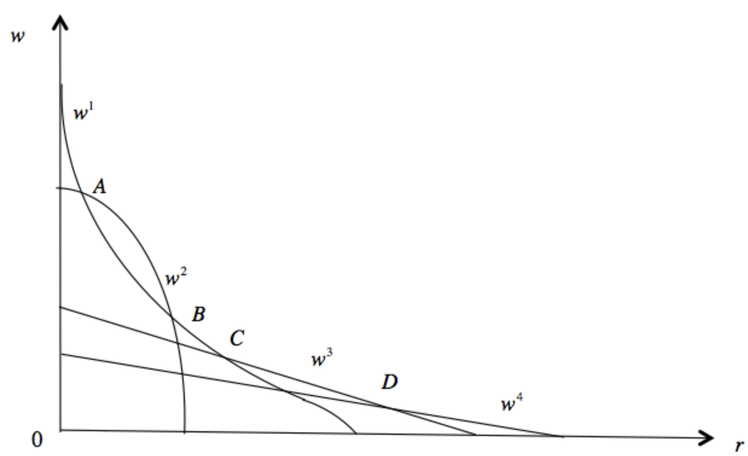


Diagram Note 4: Four wage curves, w^1 and w^2 curved, w^3 and w^4 linear. Intensity of capital falls with rising r in line with neoclassical theory at D. Paradoxical effects due to the curvatures of w^1 and w^2.

[5]Before hinting at generalisations, Samuelson worked with a two-sector model of a capital good, used in its own reproduction and for the production of a consumption good, which became the numéraire. He supposed that the capital good was, for each technique, specific for the production of the consumption good, that is, the capital good of one technique could not be used in combination with the method to produce the consumption good of another technique.

Wage curves are straight lines, if and only if either the numéraire is an eigenvector of the input-output matrix (Sraffa's standard commodity) or if relative prices are constant for all r, hence if they are equal to prices at $r = 0$, which means that they are equal to labour values. Samuelson did not make it clear that this was his assumption. It has been pointed out by Salvadori and Steedman (1988) that the intersection of the linear wage curves of two systems on the Samuelson's envelope would be dominated by combinations of the methods taken from the two systems, but Samuelson had excluded such combinations by this assumption of specificity.

[6]Reswitching has been illustrated in note 4. Reverse capital deepening occurs, if there is a third wage curve dominating the first but not the second switch in what would otherwise, without the third wage curve, be reswitching. One then has an envelope with all switches except one being of the neoclassical type, and the switch where the intensity of capital unexpectedly rises is not marked as a return of a technique that would be visible on the envelope.

[7]If a large number of straight individual wage curves (as above w^3 and w^4 in Note 5) of individual techniques are given, such that each is in part on the envelope, the envelope itself constitutes a collective wage curve for the entire spectrum of techniques. It is monotonically falling and convex. Making the transition to a continuum of techniques, one gets a smooth wage curve as envelope $\bar{w}(r)$, with $\bar{w}'(r) < 0$ and $\bar{w}''(r) > 0$. The absolute value of the slope of the tangent $-\bar{w}'(r) = k$ is the capital-intensity of the technique. If one now defines a per capita function $f(k) = \bar{w}(r) + rk(r) = \bar{w}(r) - r\bar{w}'(r)$, one finds

$$\frac{df}{dk} = \frac{df}{dr}\Big/\frac{dk}{dr} = (\bar{w}'(r) - \bar{w}'(r) - r\bar{w}''(r))/(-\bar{w}''(r)) = r$$

and

$$\frac{d^2 f}{dk^2} = -\frac{1}{\bar{w}''(r)} < 0.$$

Extending to $F(K, L) = Lf(k)$, one therefore has constructed a production function with constant returns to scale and with diminishing marginal products. The converse, the derivation of the wage curve from the production function fulfilling the marginal productivity conditions is obvious and yields $-\bar{w}'(r) = k > 0$ and $\bar{w}''(r) > 0$. (The argument is also used in the proof of Proposition 4.)

The difficulties discussed in the critique arise, if the individual wage curves are not straight. Suppose the envelope, the collective wage curve $\bar{w}(r)$, is tangent to an individual wage curve $\tilde{w}(r)$ at some $\bar{r}$. According to the construction of the production function, we must have $-\bar{w}'(\bar{r}) = -\tilde{w}'(\bar{r}) = k$, but, for the individual wage curve, k is given by $k = (1/r)/(\tilde{w}(0) - \tilde{w}(r))$, according to Diagram Note 3. The discrepancy between the two determinations of k is called declination; it is illustrated in Diagram Note 7:

Declination, the discrepancy of $k = \tan\alpha$ and $k = \tan\beta$, could be zero by coincidence, if $\tilde{w}(0)$ happened to coincide with point $P = \bar{r}\tan\alpha$, although $\tilde{w}(r)$ is not a straight line, but if we neglect the possibility of the coincidence (which has mostly been overlooked in the literature), it is clear that straight individual wage curves are not only sufficient, but also necessary for the neoclassical paradigm to hold. Wicksell effects, reswitching and reverse capital deepening are other manifestations of the problem of non-linear wage curves.

[8]See also Note 5 above.

[9]On the one hand, one has to admit that methods of production are always adapted to local conditions (institutions, geographical givens). On the other, combinations are the essential result of competition and emerge up to today in the process of globalisation. One of the most eminent and influential present-day Chinese economists writes: "Technological innovation: borrowing is the preferred option." (Lin 2012, p. 13). This refers primarily to the copying and developing of industrial processes of the more advanced countries by those, who catch up. It is typical that the technique to be copied is the dominating technique, the most advanced technique, and not one that would be less efficient and more labour-intensive. The Chinese do not copy techniques that are twenty years old, but the most recent ones, if they can, although they still live in a labour surplus economy.

[10]We recall that Samuelson postulated specificity (Samuelson 1962, p. 196), while Schumpeter postulated that technical change came about through 'new combinations' (Schumpeter 1969 [1934], pp. 12–16).

[11]If $\mathbf{A}$ is diagonalisable, with eigenvalues $\mu_1, \ldots, \mu_n$, left-hand eigenvectors (rows) $\mathbf{q}_i$; $\mathbf{q}_i\mathbf{A} = \mu_i\mathbf{q}_i$, and right-hand eigenvectors $\mathbf{x}^i$; $\mathbf{A}\mathbf{x}^i = \mu_i\mathbf{x}^i$; $i = 1, \ldots, n$; and if we write $\mathbf{l} = \mathbf{x}^1 + \cdots + \mathbf{x}^n$, where, for convenience, the $\mathbf{x}^i$ are so normalized that the coefficients in the representation of $\mathbf{l}$ as a linear combination of the $\mathbf{x}^i$ are all equal to unity, we have

$$\mathbf{p} = (\mathbf{I} - (1+r)\mathbf{A})^{-1}\mathbf{l} = w\sum_{i=1}^{n}\frac{\mathbf{x}^i}{1-(1+r)\mu_i}.$$

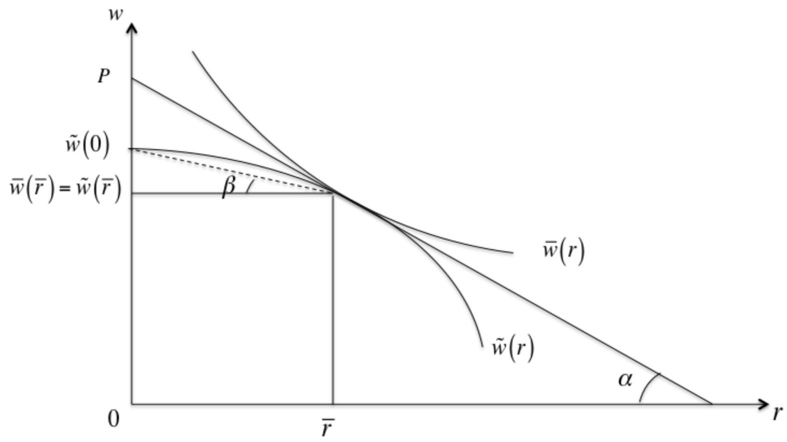


Diagram Note 7: If there is a continuum of techniques and if the wage curve of each technique is a straight line, it is tangent to the collective wage curve $\bar{w}$ as shown in the digram at $\bar{r}$. Output per head then is equal to P and the capital-labour ratio is given by $\tan\alpha$. But if the individual wage curves are not straight, like wage curve $\tilde{w}$ at $\bar{r}$, output per head is given by $\tilde{w}(0)$ and the capital-labour ratio is given by $\tan\beta$. The fact that there is a difference between P and $\tilde{w}(0)$ and between $\tan\alpha$ and $\tan\beta$ is called declination.

If $\mu_1 > 0$ is the Frobenius eigenvalue with $\mathbf{q}_1 > 0$ and $\mathbf{x}^1 > 0$, we get, with $\mathbf{q}_1$ as numéraire vector, because $\mathbf{q}_1\mathbf{x}^j = 0$; $j = 2, \ldots, n$;

$$1 = \mathbf{q}_1\mathbf{p} = \frac{w\mathbf{q}_1\mathbf{x}^1}{1-(1+r)\mu_1} = w\frac{1+R}{R-r}\mathbf{q}_1\mathbf{x}^1$$

with $\mu_1 = 1/(1+R)$; R maximum rate of profit. Hence

$$w = \frac{R-r}{(1+R)\mathbf{q}_1\mathbf{x}^1},$$

which is Sraffa's linear wage curve, except for the difference in normalisation.

12 Whether the deviations from linearity are strong or weak is a matter of judgement. See Mariolis and Tsoulfidis (2014).

13 This has been maintained consistently by Anwar Shaikh and his school (Shaikh 2016).

14 The eigenvalues of large random matrices tend to zero (Goldberg and Neumann 2003), and if we set them to zero, the formula for prices of Note 11 becomes

$$\mathbf{p} = w\Big[\frac{1+R}{R-r}\mathbf{x}^1 + \mathbf{x}^2 + \cdots + \mathbf{x}^n\Big].$$

Let net output vector $\mathbf{d}$ be written as $\mathbf{d} = \mathbf{q}_1 + \cdots + \mathbf{q}_n$ in analogy to the representation of $\mathbf{l}$ in Note 10. The wage in terms of $\mathbf{d}$ then is

$$w = \frac{1}{\mathbf{d}\hat{\mathbf{p}}} = \frac{1}{\frac{1+R}{R-r}\mathbf{q}_1\mathbf{x}^1 + \mathbf{q}_2\mathbf{x}^2 + \cdots + \mathbf{q}_n\mathbf{x}^n}$$

The wage curve is linear, if

$$\mathbf{q}_2\mathbf{x}^2 + \cdots + \mathbf{q}_n\mathbf{x}^n = (\mathbf{d} - \mathbf{q}_1)(\mathbf{l} - \mathbf{x}^1) = 0,$$

therefore if the deviations $\mathbf{m} = \mathbf{d} - \mathbf{q}_1$ of the 'Sraffa vector' $\mathbf{q}_1$ from the numéraire vector $\mathbf{d}$ and of the 'Marx vector' $\mathbf{x}^1$ from the labour vector $\mathbf{l}$, $\mathbf{v} = \mathbf{l} - \mathbf{x}^1$ are orthogonal, and this will be the case if $\mathbf{m}$ and $\mathbf{v}$ are uncorrelated, for then, $\mathbf{mv} = n\overline{m}\,\overline{v}$ and $\overline{v}$ can be shown to vanish, if $\mathbf{A}$ is random. An 'approximate' surrogate production function is proposed on the basis of these assumptions, leading to linear wage curves in Schefold (2012).

15 In between the totally disaggregated general equilibrium model of the Arrow-Debreu type and the aggregate production function models we find the early neoclassical general equilibrium models described by Garegnani (1960) and and Petri (2004), in which distribution between capital and labour is analysed by assuming that a

nominal quantity of capital, fixed in terms of the numéraire, is given as the supply. The demand for capital goods follows from the conditions of reproduction in a steady state, and the endowments permitting the maintenance of the steady state are endogenous. It can be shown that the conditions for such an equilibrium to exist and to fulfill appropriate stability conditions are very similar to those needed to prove the possibility of aggregation of a production function (Schefold 2016).

[16]Samuelson's interpretation of the maximum rate of profit as the maximum rate of growth is, of course, formally correct but the comparisons here made refer to stationary systems, as he himself pointed out. We could therefore also contrast the productivity of labour with the productivity of capital, for the maximum rate of profit gives profit divided by capital, when profit absorbs the entire product. (The maximum rate of profit can by the way also be seen as the inverse capital-output ratio, with capital goods valued at the prices pertaining to this maximum rate.)

[17]Strictly speaking, m measures the number of grid lines counted along FG and FD respectively, so that m should more rigorously be defined as the largest natural number smaller than $\sqrt{\gamma s}$.

[18]Note that the square root here imposes itself. For if one lets m grow with s, using $m = s^{\alpha}$, it is clear that $\alpha > 0$ (otherwise no growth) and $\alpha < 1$ (otherwise m overtakes s). The convergence of

$$\exp\Big[\frac{m(m+1)}{2(s-m)}\Big] = \exp\frac{s^{2\alpha}+s^{\alpha}}{2s-2s^{\alpha}} = \exp\frac{1+s^{-\alpha}}{2s^{1-2\alpha}-2s^{-\alpha}}$$

then requires $\alpha = 1/2$.

[19]We visualize the distribution for the main example. The techniques used in 100 sectors by 10 countries can be identified by means of a number with 100 digits. Each digit expresses the country of origin of the method of production used in the industry, expressed by the position of the digit in this number. Accordingly, in order to keep the techniques distinct in the grid, outputs per head and maximum rates of profit must also be measured in terms of at least 100-digit numbers, even if this degree of precision is far beyond that usually employed in economics. That the techniques, either as given by the input-output coefficients or as represented by the wage curves, are actually distinct can be proven by means of the theory of regular systems (Schefold, 1971). To insist on the distinctness is important in order to model the difference between the understanding of techniques as separate activities from the set-theoretical approach which was promoted in the Arrow-Debreu version of general equilibrium theory in the 1950s and beyond. It is plausible that the techniques will cluster in the middle between the extremes, without, however, getting really sparse at the ends. To assume an independent probability distribution for the anchor points on the lines and columns of the grid is not in contradiction with the idea of a normal distribution on the margin of the grid. An analogy may help to visualise this. The individuals in a population of adults can be characterised by their height and their intelligence, with nearly normal distributions. There will be a clustering around the average height, say 175 cm, and around the average intelligence index, say 100. Assume that all combinations of height and intelligence are equally probable in this grid. The probability distribution, although it is uniform, will nonetheless express a clustering indirectly: Since there are more individuals, who are of average intelligence, as measured on the abscissa (scaled by the index), the outcome for a person who is tall, as measured on the ordinate (scaled in cm), is that he/she will also more likely be of average intelligence. It basically means that height and intelligence are not correlated.

[20]Putting $w^{s}(0) = 0$ for simplicity.

[21]Anwar Shaikh (1987) has shown that this growth process with constant shares and a rising capital-intensity creates the illusion of a Cobb-Douglas production function.

[22]The bivariate uniform distributions follow Ferguson (1995). The expected value for the normal distribution is in all numerical experiments 7 and the variance 1.

[23]Each country thus is treated as if it was autarchic. Advantages of international trade then are blurred. The assumption is possible only, if imported "goods", at the given level of aggregation, are also produced domestically.

[24]Which anchor points get on the envelope, apart from those on the upper right of the convex hull of the anchor points, may be inferred from Diagram 6. The coordinates of the anchor points of the constellation represented in the Diagram are given by (c, b) and (d, a). Consider the dashed line. If it is turned in such a way that it keeps crossing the intersection of the constellation, denominated by (i, j) and if $c < r_k < d$ and $a < k < b$, it is clear that the anchor point (r_k, k) will represent a line (the dashed line) that just touches the envelope, and straight lines parallel to and above the dashed line will be on the envelope. If the dashed line moves between the limits given by the constellation, the anchor point will move on a hyperbola $k = [j/(i - r_k)]r_k$ in function of r_k. The asymptotes of this hyperbola will be given by a horizontal and a vertical line running through (i, j) and the hyperbola itself will go through the anchor points of the constellation. The hyperbola therefore will separate the area below and to the left of the anchor points of the constellation into two parts, of anchor points that stand for lines of shortcuts below or on the envelope.

References

Aghion, Philippe and Howitt, Peter (1998): *Endogenous Growth Theory.* Massachusetts Institute of Technology.

Bliss, Christopher J. (1970): Comment on Garegnani. *Review of Economic Studies*, vol. 17, pp. 437–438.

Ferguson, Thomas S. (1995): A class of symmetric bivariate uniform distributions. *Statistical Papers*, vol. 36 (1), p. 31.

Garegnani, Pierangelo (1960): *Il Capitale nelle Teorie della Distribuzione.* Milan: Giuffrè.

Garegnani, Pierangelo (1970): Heterogeneous Capital, the Production Function and the Theory of Distribution, *Review of Economic Studies*, vol. 37 (3), pp. 407–436.

Garegnani, Pierangelo (1976): On a Change in the Notion of Equilibrium in Recent Work on Value and Distribution, in: Murray Brown, Kazuo Sato, and Paul Zarembka (eds.), *Essays on Modern Capital Theory.* Amsterdam: North-Holland Publication, pp. 25–45.

Goldberg, G. and Neumann, M. (2003): Distribution of Subdominant Eigenvalues of Matrices with Random Rows. *Society for Industrial and Applied Mathematics Journal on Matrix Analysis and Application* 24 (3), pp. 747–761.

Hagemann, Harald (2020): The Cambridge-Cambridge Controversy on the Theory of Capital: 50 Years After. Introduction. *European Journal of Economics and Economic Policies*, vol. 17 (2), pp. 196–207.

Han, Zonghie (2003): *"Paradoxa" in der Kapitaltheorie: eine empirische Untersuchung der Reverse Capital Deepening- und Reswitching-Phänomene anhand der linearen Programmierung im Rahmen der Kapitalkontroverse.* Marburg: Metropolis.

Han, Zonghie and Schefold, Bertram (2006): An Empirical Investigation of Paradoxes: Reswitching and Reverse Capital Deepening in Capital Theory. *Cambridge Journal of Economics*, vol. 30 (5), pp. 737–765.

Harcourt, Geoffrey C. (1972):*Some Cambridge Controversies in the Theory of Capital. Cambridge*: Cambridge University Press.

Kurz, Heinz; Salvadori, Neri (1999): *Theory of Production. A Long-Period Analysis.* Cambridge: Cambridge University Press.

Kurz, Heinz, ed. (2000): *Critical Essays on Piero Sraffa's Legacy in Economics.* Cambridge: Cambridge University Press.

Kurz, Heinz (2020): The Theory of Value and Distribution and the Problem of Capital. *European Journal of Economics and Economic Policies*, vol. 17 (2), pp. 241–264.

Levhari, David (1965): A Nonsubstitution Theorem and Switching of Techniques. *The Quarterly Journal of Economics*, vol. 79 (1), pp. 98–105.

Lin, Justin Yifu (2012): *Demystifying the Chinese Economy.* Cambridge: Cambridge University Press.

Mainwaring, Lynn and Steedman, Ian (2000): On the Probability of Reswitching and Capital Reversing in a Two-sector Sraffian Model, in: Kurz, H. (ed.) (2000): *Critical Essays on Piero Sraffa's Legacy in Economics*. Cambridge: Cambridge University Press, pp. 323–354.

Mariolis, Theodore and Tsoulfidis, Lefteris (2014): On Bródy's Conjecture: Theory, Facts and Figures about Instability of the US Economy. *Economic Systems Research*, 26 (2), pp. 209–223.

Pasinetti, Luigi L. (1966): Changes in the Rate of Profits and Switches of Techniques. *The Quarterly Journal of Economics*, vol. 80 (4), pp. 503–517.

Petri, Fabio (2004): *General Equilibrium, Capital and Macroeconomics: A Key to Recent Controversies in Equilibrium Theory*. Cheltenham, UK and Northampton, MA: Edward Elgar Publishing.

Petri, Fabio (2011): On the Likelihood and Relevance of Reswitching and Reverse Capital Deepening, in: Salvadori, N., Gehrke, C. (eds.), *Keynes, Sraffa and the Criticism of Neoclassical Theory: Essays in Honour of Heinz Kurz*. London and New York: Routledge, pp. 380–418.

Robinson, Joan (1953–1954): The Production Function and the Theory of Capital. *Review of Economic Studies*, vol. 21, pp. 312–320.

Salvadori, Neri and Steedman, Ian (1988): No Reswitching? No Switching. *Cambridge Journal of Economics*, 12 (4), pp. 481–486.

Samuelson, Paul (1962): Parable and Realism in Capital Theory: The Surrogate Production Function. *Review of Economic Studies*, 1962, vol. 29 (3), pp. 193–206.

Samuelson, Paul (1966): A Summing Up. *The Quarterly Journal of Economics*, vol. 80 (4), pp. 568–583.

Samuelson, Paul (1966 [1961]): Economists and the History of Ideas (Presidential Address, A.E.A. 1961), in: Joseph E. Stiglitz, (ed.), *The Collected Scientific Papers of Paul A. Samuelson*, Massachusetts Institute of Technology, pp. 1499–1516.

Schefold, Bertram (1971): Piero Sraffas Theorie der Kuppelproduktion, des fixen Kapitals und der Rente (Mr. Sraffa on Joint Production). Private Print. Online as: `https://www.wiwi.uni-frankfurt.de/fileadmin/user_upload/dateien_abteilungen/abt_ewf/Economic_Theory/publications/Theorie_der_Kuppelproduktion_Bertram_Schefold_1971.pdf`.

Schefold, Bertram (1976): Different Forms of Technical Progress, *The Economic Journal*, 86, pp. 806–819.

Schefold, Bertram (2008): Savings, Investment and Capital in a System of General Intertemporal Equilibrium: An Extended Comment on Garegnani with a Note on Parrinello, in: Guglielmo Chiodi and Leonardo Ditta (eds.), *Sraffa or an Alternative Economics*. Basingstoke: Macmillan, pp. 127–184.

Schefold, Bertram (2013a): Approximate Surrogate Production Functions, *Cambridge Journal of Economics*, 37 (5), pp. 947–983.

Schefold, Bertram (2013b): Only a Few Techniques Matter! On the Number of Curves on the Wage Frontier, in: Levrero, E., Palumbo, A. and Stirati, A. (eds.), *Sraffa and the Reconstruction of Economic Theory: Volume One*, Theories of Value and Distribution. Basingstoke, UK: Palgrave Macmillan, pp. 46–69.

Schefold, Bertram (2016): Marx, the Production Function and the Old Neoclassical Equilibrium: Work-able under the Same Assumptions? With an Appendix on the Likelihood of Reswitching and Wicksell Effects. *Centro Sraffa Working Papers*, No 18, March. Online as *Centro Sraffa Working Papers*, ISSN 2284–2845.

Schefold, Bertram (2017): Pluralism in Economics: The Case of the Cambridge School. Proposal for a Transformation of the Critique of Capital, in: Pieter Hennipman Stichting (ed.) *Het Gesproken Woord*. Pieter Hennipman Diner 2017. Amsterdam, pp. 16–46.

Schefold, Bertram (2018): The Improbability of Reswitching, the Certainty of Wicksell-Effects and the Poverty of Production Functions: The Cambridge Critique of Capital Transformed, in: Luca Fiorito, Scott Scheall and Carlos E.Suprinyak, (eds.). *Research in the History of Economic Thought and Methodology*, vol. 35B, including a Symposium on New Directions in Sraffa Scholarship. Bingley, UK: Emerald Publishing Limited 2018, pp.171–194.

Schefold, Bertram (2020): What Remains of the Cambridge Critique of Capital Theory, if Reswitching and Reverse Capital Deepening are Empirically Rare and Theoretically Unlikely? *European Journal of Economics and Economic Policies*, vol. 17 (2), pp. 220–240.

Schefold, Bertram (2021): Transformations of the Cambridge Critique. *Indian Economic Journal*. To be published.

Schumpeter, Joseph Alois (1969 [1934]): *The Theory of Economic Development*, transl. by Redvers Opie, Reprint. London: Oxford University Press.

Shaikh, Anwar (1987): Humbug Production Function. *The New Palgrave. A Dictionary of Economics*, vol. 2. London: Macmillan, pp. 690–691.

Shaikh, Anwar (2016): Capitalism: *Competition, Conflict, Crises*. Oxford: Oxford University Press.

Shaikh, Anwar; José Alejandro Coronado; Luiza Nassif-Pires (2020): On the Statistical Regularities of Prices of Production. *European Journal of Economics and Economic Policies*, vol. 17 (2), pp. 265–275.

Shapiro, Samuel S. and Wilk, Martin B. (1965): An analysis of variance test for normality (complete samples). *Biometrika*, vol. 52 (3/4), pp. 591–611.

Sraffa, Piero (1960): *Production of Commodities by Means of Commodities*. Cambridge: Cambridge University Press.

Timmer, Marcel P.; Erik Dietzenbacher; Bart Los; Robert Stehrer; Gaaitzen J. De Vries (2015): An illustrated user guide to the world input–output database: the case of global automotive production. *Review of International Economics*, vol. 23 (3), pp. 575–605.

Timmer, Marcel P.; Bart Los; Robert Stehrer; Gaaitzen J. De Vries (2016): An anatomy of the global trade slowdown based on the WIOD 2016 release. Technical report, Groningen Growth and Development Centre, University of Groningen.

Weizsäcker, Carl Christian von (1971): *Steady State Capital Theory*. (Lecture Notes in Operations Research and Mathematical Systems; 54). Berlin: Springer.

Weizsäcker, Carl Christian von (2020): Böhm-Bawerk and Hicks Modernized. *European Journal of Economics and Economic Policies*, vol. 17 (2), pp. 208–219.

Zambelli, Stefano; Thomas Fredholm; Ragupathy Venkatachalam (2017): Robust Measurement of National Technological Progress. *Structural Change and Economic Dynamics*, vol. 42, pp. 38–55.

Zambelli, Stefano (2018): The aggregate production function is NOT neoclassical. *Cambridge Journal of Economics*, 42, 2018, pp. 383–426.