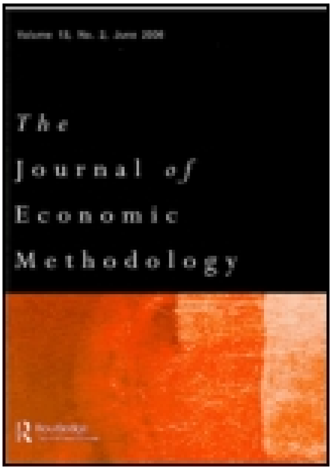

Journal of Economic Methodology
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/rjec20

Unobservability, tractability and the battle of assumptions

Frank A. Hindriks [a]
[a] Erasmus University Rotterdam E-mail:
Published online: 12 Apr 2011.

Journal of Economic Methodology 12:3, 383–406 September 2005 Routledge Taylor & Francis Group

Unobservability, tractability and the battle of assumptions

Frank A. Hindriks

Abstract Economic models often include unrealistic assumptions. This does not mean, however, that economists lack a concern for the truth of their assumptions. Unrealistic assumptions are frequently imposed because the effects are taken to be negligible or because the problem at hand is intractable without them. Using the Musgrave-Mäki typology as the point of departure, these claims are defended with respect to theories proposed by Solow, Hall and Roeger concerning productivity growth and the mark-up. Since they are unobservable, their values need to be inferred from the values of observable variables. Assumptions such as perfect competition and constant returns to scale are used for making this inference or measurement problem tractable. Other assumptions are justified in terms of negligibility. These findings support the fecundity of the (amended) Musgrave-Mäki typology of assumptions – including the notion of a tractability assumption proposed here. Finding ways of relaxing tractability assumptions turns out to be an important source of progress in economics.

Keywords: negligibility assumption, tractability assumption, Musgrave, F-twist, unrealistic assumptions, unobservables, measurement, progress

1 INTRODUCTION

Scientific theories are bound to be unrealistic in one way or another, notably by containing unrealistic assumptions. Musgrave (1981) recognized that such assumptions are often imposed for good reasons. He argued, for example, that scientists frequently abstract from certain factors because they regard their effects as negligible. He proposed to call the assumptions scientists use for this purpose ‘negligibility assumptions’. Musgrave proposed a typology of assumptions that classifies them according to the role they play in the theory of which they are part. In addition to negligibility assumptions, he distinguished domain assumptions and heuristic assumptions.

Mäki (2000) proposed several refinements and reformulations of the typology. In this paper, I propose further amendments. In particular, I introduce the notion of a tractability assumption. The underlying idea is that certain assumptions are imposed in order to make particular problems tractable. What I propose to call ‘the amended Musgrave-Mäki typology’

Journal of Economic Methodology ISSN 1350-178X print/ISSN 1469-9427 online

DOI: 10.1080/13501780500223619

will be applied to a series of economic theories. In particular, Solow's (1957) method for measuring productivity growth and the methods Hall (1988) and Roeger (1995) developed for measuring the mark-up of price over marginal cost – both of which have been derived from Solow's method – will be investigated. This case study identifies several tractability and negligibility assumptions and thereby confirms the fecundity of the amended Musgrave-Mäki typology.

I also investigate why Solow, Hall and Roeger use tractability assumptions. I argue that they are imposed because important variables of the relevant models are unobservable. The notions of marginal cost and the price of capital as used by economists are theoretical constructs and do not correspond to data provided by statistical agencies, except perhaps in exceptional circumstances. Roughly speaking, the tractability assumptions in this case – for instance, the assumptions of perfect competition and constant returns to scale – are statements to the effect that such circumstances obtain, even though this may in fact not be the case. I go on to argue that many economists regard relaxing tractability assumptions in order to arrive at more realistic models as an important way of achieving progress.

The picture that emerges is very different from other characterizations that are prominent in economic methodology. It is often argued that economists are not very concerned about the real world. They are characterized as scientists who play in a sandbox or as instrumentalists that only have predictive success as their goal. The case considered here, however, reveals a deep concern for realisticness. Thomas Mayer claims that 'when it comes to choosing hypotheses economists often pay much attention to the realisticness of assumptions' (1999: 319). He also claims that it is hard to document this. The case to be examined provides such documentation.

Section 2 discusses the measurement methods of Solow, Hall, and Roeger and presents the four core assumptions of the case under investigation. Those who are familiar with these methods may wish to skip this section. Section 3 classifies these core assumptions in terms of what I call the amended Musgrave-Mäki typology. Furthermore, it discusses the role of unobservability in the imposition of false assumptions. The main analyses of the paper – those concerning unobservability, tractability, and progress – can be found in section 4.

2 THE MEASUREMENT OF PRODUCTIVITY GROWTH AND THE MARK-UP

Before discussing the economics of these methods, let me remark on their position within economics. The three methods are typically applied at the level of industries, and, as such, they are part of industrial organization. The methods of Hall and Roeger belong to the *New Empirical Industrial Organization* (*NEIO*; see Bresnahan 1989 for a discussion). *NEIO*-studies are studies of market power that take marginal cost to be unobservable. The data used are time series of prices and quantities of inputs and output, as well as variables that determine demand. *NEIO* differs from the *Structure-Conduct-Performance* (*SCP*) paradigm in that it focuses on conduct and performance of individual industries rather than on cross-section studies of the relationship between structure, conduct, and performance across industries (see Church and Ware 2000). The three methods are relevant for macroeconomics as well. Apart from the fact that productivity growth can be measured at the level of entire economies, productivity growth plays an important role in several theories of the business cycle.

2.1 Solow's method for measuring productivity growth

Robert M. Solow (1957) proposed a measure of productivity growth that has become known as the Solow residual (SR). Solow started from the idea that the relation between the inputs and the output of a production process can be represented by a production function. The basic insight underlying his method is that a production function can be used for this purpose not only at the level of an individual firm but also at an aggregate level such as that of an industry or even an entire economy. Productivity growth can then be determined indirectly by studying the developments of the way in which the various inputs contribute to the output.

Solow's point of departure was a general production function, a production function about which no parametric assumptions are made. He showed that productivity growth equals the growth in output minus the sum of the weighted growth of the input factors. Because his method does not attribute productivity growth to particular production factors, the SR became also known as total factor productivity (TFP) growth. Equation (1) gives the expression of the Solow residual with the cost shares of the input factors functioning as the weights of their growth:

$$SR = \frac{\Delta Q}{Q} - \sum \alpha_j^C \frac{\Delta W_j}{W_j} \tag{1}$$

where Q is the quantity of the product; W_j is the price of input j with $j=\{N, K\}$; N is labour and K is capital; Δ designates the change in a variable over time; $\alpha_j^C = \frac{W_j X_j}{MC \cdot Q}$ is the cost share of input j; X_j is the quantity of input j; and MC is marginal cost (as is often done, I abstract from intermediate inputs).

The expression that Solow actually used for calculating productivity growth was formulated in terms of income shares of inputs rather than cost shares as weights. Solow imposed two assumption in order to derive that expression: perfect competition [A1: PC], and constant returns to scale [A2: CRS]. In perfect competition, price equals marginal cost. Hence, the first assumption can be expressed in terms of the mark-up, $\mu = \frac{P}{MC}$ (with P as price), which equals one if competition is perfect:

$$\mu = 1 \quad [\mathrm{A1:PC}] \tag{2}$$

This implies that the cost shares of the inputs, α_j^C, are equal to the income shares, $\alpha_j^p = \frac{W_j X_j}{PQ}$:

$$\alpha_j^C = \alpha_j^P \tag{3}$$

Because of this equality, income shares can replace cost shares in equation (1).

The second assumption is that of constant returns to scale (CRS). In mathematical terms, the returns to scale are constant if the cost shares adds up to 1:

$$\sum \alpha_j^C = 1 \ [\mathrm{A2:CRS}] \tag{4}$$

Using the assumption of constant returns, one of the cost shares can be eliminated. Since cost shares are equal to income shares given the assumption of perfect competition, one of the income shares can be eliminated. Usually it is the capital income share that is eliminated. It can be eliminated using the following expression:

$$\alpha_K^P = 1 - \alpha_N^P \tag{5}$$

Equation (5) is derived from equations (3) and (4). Hence, it presupposes both the assumption of perfect competition and that of constant returns to scale. The traditional formula for the Solow residual can now be derived in a straightforward way. First, reformulate equation (1) in terms of income shares – using equation (3) and thereby the assumption of perfect competition. Second, eliminate the capital income share from the resulting equation – using equation (5) and hence the assumption of constant returns to scale. These two steps lead to equation (6):

$$SR = \Delta q - \Delta k - \alpha_N^P (\Delta n - \Delta k) \tag{6}$$

where Δq represents the log difference of Q and is an approximation of the percentage growth. This can be rewritten as follows:

$$SR = \Delta q' - \alpha_N^P \Delta n' \tag{6'}$$

where $\Delta q'$ designates the log difference of Q/K and $\Delta n'$ that of N/K.

2.2 Hall and Roeger: measuring the mark-up

By the time Hall and Roeger wrote their articles, two features of US productivity growth as measured by Solow's method had been discovered.

First, the SR is procyclical: it increased in booms and decreased in recessions. Second, the dual SR differs from the primal SR in practice even though they are equivalent in theory. The discovery of the second feature required a theoretical innovation. Hulten (1986) showed that a measure of productivity growth similar to the original SR can be derived from a cost function rather than a production function using the same assumptions as Solow had used. Because of the use of duality theory, this second measure of total factor productivity growth became known as the dual SR (the original SR became known as the primal SR).[1]

Hall offered an explanation of the first feature, and Roeger offered an explanation of the second feature. Both explain the respective features in terms of a failure of the assumption of perfect competition. This implies that the two features are not genuine features of productivity growth. Instead, they are artefacts of the assumption of perfect competition that is used in the measurement process. In Hall's case, this implies that real productivity growth may not be as procyclical as it was thought to be, since the procyclicality of measured productivity growth is (at least partly) caused by the mark-up component that had not been taken into account. In line with this, Hall claimed that '[i]n the presence of market power, measured total factor productivity (based on the assumption of competition) is spuriously procyclical' (1987: 424). In Roeger's case, it implies that the lack of correlation between the primal and the dual residual is for the most part only apparent (1995: 328). When the mark-up component is taken into account, more than 90 per cent of the difference is explained (1995: 325).

Both Hall and Roeger support their claim that competition is often imperfect with evidence. They do this by deriving a method for measuring the mark-up of price over marginal cost from Solow's method of measuring productivity growth. The relation between Solow's method and that of Hall is as follows: whereas Solow made an assumption concerning the mark-up [A1: PC] and designs a method for measuring productivity growth, Hall makes an assumption concerning productivity growth and develops a way of estimating the mark-up. The first step made by Hall, then, is relaxing the assumption of perfect competition [A1: PC]. Without the assumption of perfect competition, the SR does not anymore represent productivity growth only but has both a productivity growth and a mark-up component.

In order to distinguish between SR and real productivity growth, the latter will be designated by τ from now on. Hall goes on to introduce an assumption concerning productivity growth. He assumes that productivity growth is constant over time, deviations being randomly distributed [A3: CPG]:

$$\tau = \theta + \varepsilon \text{ [A3 : CPG]} \tag{7}$$

where θ is constant productivity growth, and ε is an error term. This is the third assumption that plays a central role in the cases under investigation.

Once the assumption of perfect competition, [A1: PC], has been relaxed, the mark-up appears in the estimation equation. Hall assumes that the mark-up μ remains constant over the business cycle, the fourth core assumption of the case under investigation [A4: CM]:

$$\mu = c \quad [\mathrm{A4:CM}] \tag{8}$$

This means that the mark-up does not have a random component. Given this assumption, the mark-up has no influence on the error term of the equation. Thus, the error term of the estimation equation is identical to the random component of technical change, ε. As we will see below, this is very important for Hall's method.

Hall's estimation equation can now be derived as follows. First, replace SR in equation (1) by τ. Second, eliminate the capital cost share using equation (4). Third, reformulate the equation in terms of income shares. Thus, the mark-up appears in the equation, as competition is no longer assumed to be perfect. By reformulating the resulting equation in terms of log differences and rearranging terms equation (9) results:

$$\Delta q - \Delta k = \mu \cdot \alpha_N^P (\Delta n - \Delta k) + \theta + \varepsilon \tag{9}$$

This in turn can be rewritten as follows (see Hall 1986: 291, 1988: 926; recall that $\Delta q'$ designates the log difference of Q/K.):

$$\Delta q' = \mu \cdot \alpha_N^P \Delta n' + \theta + \varepsilon \tag{9$'$}$$

Just as in equation (6′) the income share of capital has been eliminated. Let us pause a moment and recapitulate how this is done. Solow could eliminate one of the income shares because he assumed both constant returns to scale and perfect competition. Hall and Roeger, however, do not assume perfect competition. They can eliminate the capital income share nevertheless because they include the mark-up of price over marginal cost in their equation. A cost share is by definition equal to the product of the mark-up and the relevant income share: $\alpha_j^C = \mu \cdot \alpha_j^P$. Since constant returns to scale means that the cost shares add up to one, this assumption allows one to eliminate the capital income share in virtue of the following equation: $\alpha_K^P = \frac{1}{\mu} - \alpha_N^P$. This means that eliminating marginal cost from the estimation equation as is done by Solow is not necessary for eliminating one of the income shares. In other words, even without the assumption of perfect competition, the assumption of constant returns to scale can be used for eliminating one of the income shares.

In order to get to the heart of Hall's method, we need to discuss further details including the estimation method he used. The SR as measured by equation (6′) has turned out to be procyclical, as has been confirmed by numerous investigations. This makes it problematic to estimate equation (9′) using a simple estimation method such as ordinary least squares. The error term is likely to be correlated to cyclical variables such as output and employment. If there is such a correlation, ordinary least squares estimates will be inconsistent, which is to say that the estimator of the mark-up does not converge on the true parameter. Because of this, the method of ordinary least squares is no longer the appropriate estimation method. The inconsistency of the estimator can be resolved by using instrumental variable estimation. Instrumental variables should meet two conditions: first, they should be correlated with the independent variable; second, they should not be correlated with the error term. In a sense, instrumental variables are proxies for the real variables (even though the real variables still play a role in the estimation process). Given the assumption that productivity growth is not cyclical but constant, [A3: CPG], the conditions that an instrument has to meet translate into the following conditions for the case at hand: (1) the instrument should be correlated with employment; (2) the instrument should not be correlated with the error term of productivity growth.

Hall formulates these conditions in a slightly different way (1986: 932, 1987: 443, 1988: 924, 932). With regard to the first condition, he claims that an instrument should be an important outside, i.e. exogenous, determinant of both output and employment. The fact that he mentions output in addition to employment makes sense given the procyclicality of the SR, i.e. productivity growth as measured by equation (6). Equation (6) implies that if the SR and employment are cyclical, output is cyclical as well. The requirement that the instrument should be exogenous ensures that it does not introduce a new source of correlation between the error term and the independent variable. With regard to the second condition, Hall claims that the instrument should not cause shifts in productivity and that it should not be influenced by variables that do. These requirements are implied by condition (2) as formulated by me, given the assumption of constant productivity growth, [A3: CPG]. If the instrument were to cause shifts in productivity growth or if it were influenced by variables that do, it would be correlated with the error term of productivity growth. The two conditions are alternatively referred to as 'identifying hypotheses' or 'identifying assumptions' (Hall 1988: 926 and Roeger 1995: 328 respectively).

We are now in a position to understand the crux of Hall's method. He describes it as follows: 'My approach, stripped to its absolute basics, is to choose as the estimate of μ the value that is just high enough to leave the residual uncorrelated to the business cycle' (Hall 1986: 291, see also p. 295). As will be clear from the preceding discussion, the use of the instrumental variable estimation method ensures that the residual is uncorrelated to the business cycle. The cyclicality of SR is thus attributed completely to the mark-up rather than to productivity growth. The consequence of this is that SR as measured by equation (6) is an inadequate measure of productivity growth. It is biased because apart from productivity growth it includes a mark-up component. This in turn implies that Hall estimates the mark-up in such a way that it fully explains the cyclicality of productivity growth as measured by equation (6′).

Let us now turn to Roeger's method for measuring the mark-up. Roeger builds on Hall's method. He derives an estimation equation in which productivity growth does not appear anymore. Because of this, he does not need to make an assumption concerning productivity growth, which means that [A3: CPG] can be relaxed. Consequently, Roeger does not have to use instrumental variables. Roeger claims that his method is superior to Hall's. As we will see below, the fact that he does not have to use instrumental variables plays a key role in Roeger's claim concerning the superiority of his method. Roeger can eliminate productivity growth from the estimation equation because in addition to the primal Solow residual he uses the dual Solow residual (see the beginning of this section). By subtracting the dual residual from the primal one, productivity growth drops out of the equation. After rearranging the terms, an expression follows that can be used to estimate the mark-up using ordinary least squares (see the appendix for a derivation):

$$SR - SRP = L\Delta x + \omega \tag{10}$$

where SRP is the dual or price-based Solow residual; ω is an error term; L is the Lerner index – an alternative notation of the mark-up – and Δx is a composite variable given by $\Delta x = (\Delta q - \Delta k) + (\Delta p - \Delta w_K)$, which is equivalent to $\Delta x' = \Delta pq - \Delta w_K k$. The Lerner index is defined as $L = \frac{P - MC}{P}$; the relationship between the Lerner index L and the estimation parameter used by Hall, μ, is given by: $L = 1 - \frac{1}{\mu}$.

Roeger maintains the assumption of constant returns to scale, [A2: CRS]. Just as Hall, he uses it for eliminating the income share of capital from the estimation equation. Because of this, he does not need to use data on the level of capital cost directly. However, he needs data on the change of capital cost (as is implied by the definition of $\Delta x'$). Their calculation requires data on the level of capital cost. Therefore, even though they do not play a role directly as input of the estimation equation, they play an indirect role as the basis for the data on changes in capital cost, which are used directly for estimating equation (10). Roeger also maintains the assumption of a constant mark-up [A4: CM].

In sum, Solow proposed a method for measuring productivity growth, assuming perfect competition, [A1: PC], and constant returns to scale, [A2: CRS]. Starting from Solow's method, Hall developed a method for measuring the mark-up, relaxing the assumption of perfect competition, [A1: PC]. Hall's method assumes constant returns to scale, [A2: CRS], constant productivity growth, [A3: CPG], and a constant mark-up, [A4: CM]. Roeger builds on Hall's method, and proposes a version that does not require an assumption about the development of productivity growth: he relaxes [A3: CPG]. Because of this, Roeger's method does not require the use of instrumental variables, whereas Hall's does. Roeger's method only assumes constant returns to scale, [A2: CRS], and a constant mark-up, [A4: CM].

3 UNOBSERVABILITY AND KINDS OF ASSUMPTIONS

Why do Solow, Hall and Roeger use the assumptions they impose? I will argue that their main motivation derives from the unobservability of some of the economic variables relevant to the variable they want to measure. In addition to this, I will argue that the assumptions can be justified in terms of negligibility and tractability and I will show that in several cases the economists mentioned justify the assumptions in such terms themselves. Related to this, I will argue that the case under investigation provides support for the amended Musgrave-Mäki typology of assumptions that distinguishes among others negligibility and tractability assumptions. This section focuses on the considerations Solow, Hall, and Roeger provide as reasons for imposing the assumptions discussed in the previous section and provides a diagnosis of them. In the next section, the analysis of the reasons for imposing assumptions will be used for providing an account of the conception of progress that is implicit in all this. Let me start by introducing the amended Musgrave-Mäki typology of assumptions.

In presenting the Musgrave-Mäki typology, I will mostly rely on Mäki (2000). Mäki revealed several shortcomings of Musgrave's proposal and suggested various changes. I regard most of these changes as improvements. In spite of this, I propose two further amendments. First, the assumptions of this typology should be seen as second-order assumptions that provide reasons for imposing first-order assumptions that in turn are used for building models. Although this fits with Musgrave (1981) and hints in this direction can be found in Mäki (2000), neither of them has made this claim. Second, the notion of a tractability assumption should replace Musgrave's category of heuristic assumptions and Mäki's class of early-step assumptions. I have defended these amendments in Hindriks (2005). Here, they will prove their usefulness in the application of the typology to the case study. Given this strategy I will not pause to discuss heuristic and early-step assumptions.

Most first-order assumptions with which the typology is concerned are statements according to which a certain factor F has no effect on the phenomenon under investigation. A negligibility assumption hypothesizes that such a first-order assumption is imposed because factor F has a negligible effect on the phenomenon under investigation given the purposes of the investigator(s). The second type of assumptions, that of an applicability assumption, puts forward the claim that a theory T applies only if the first-order assumption is true or if factor F exerts only a negligible influence. Finally, a tractability assumption conjectures that problem P can be made (more) tractable by imposing the first-order assumption. In case of negligibility and tractability assumptions the related first-order assumptions are (believed to be) false. This also holds for applicability assumptions according to which factor F has a negligible influence rather than no influence at all. The first-order assumption associated with applicability assumptions is only (supposed to be) true in a subset of cases.

The notion of a tractability assumption requires some more elucidation. On my view, tractability is a matter of solubility or of the efficiency of a solution. A problem is intractable if it cannot be solved; a problem is more tractable with a certain assumption than without it if it can be solved more easily or efficiently in that case. A tractability assumption can be theoretical or empirical. Problem P may be the problem of how to apply a certain theory T to a particular situation. An assumption that is imposed in order to solve such a problem is an empirical tractability problem. An assumption that is needed for solving a problem that is independent from the application of the theory is a theoretical tractability assumption. Of course, there can be boundary cases.

3.1 Perfect competition

We are now ready to discuss the four assumptions, [A1] – [A4], in turn and to consider the role they played for Solow, Hall, and Roeger respectively. The assumption of perfect competition, [A1: PC], only plays a substantial role in Solow's method for measuring productivity growth. Solow does not comment on why he imposes the assumption of perfect competition. There is, however, an obvious reason for doing so. Marginal cost is unobservable and data on this variable are needed for calculating the cost shares of the inputs (lets for now consider a variable as unobservable if data on it are unavailable; I will discuss the notion of unobservability further below). Hall makes the following remark about an equation that is fairly similar to equation (1): 'All the variables in the […] formula are observed directly except for marginal cost and the rate of technical progress' (1986: 289). One equation with two unknown variables cannot be solved. The equation can be solved when competition is assumed to be perfect. On that assumption, marginal cost can be eliminated in favour of price, which means that cost shares can be replaced by income shares (see equation (3)). As price is readily observable, only one unknown remains and it can now be computed from the resulting equation (6′). All this means that the assumption of perfect competition is imposed in order to be able to apply the theory to reality, in other words to make the measurement of productivity growth tractable. This in turn implies that Solow's assumption of perfect competition is an empirical tractability assumption.

One might object to this interpretation that Solow may have believed that competition was in fact perfect. This would mean that he regarded the assumption as true and that a justification in terms of tractability would be redundant. There is reason to believe that this objection is invalid. Solow does not only assume perfect competition on the market for outputs. In addition to this, he assumes perfect competition on the market for inputs such as labour and capital (because of this the price of these inputs equals their marginal products; hence, W_j can figure in equation 1 rather than the unobservable marginal product). He considers it to be a price of his method that he has to make this assumption (Solow 1957: 312). This means that he would rather not impose the assumption, but only does so because he cannot do without it. This in turn fits perfectly with my interpretation of the assumption of perfect competition as an empirical tractability assumption.[2]

3.2 Constant returns to scale

The second assumption made by Solow is constant returns to scale, [A2: CRS]. According to this assumption, the cost shares of the inputs add up to one (see equation (4)). As we saw the assumption can be used to eliminate one of the cost shares. Because of the assumption of perfect competition, this in turn means that one of the income shares can be eliminated (equation (5)). Why would one want to do that? The ulterior motive is to eliminate (the level of) the cost of one of the inputs, labour or capital from the estimation equation. The cost of the inputs occurs in the model as the numerators of the cost and income shares of the inputs. By eliminating one of the income shares, (the level of) the cost of one of the inputs is eliminated.

Hall and Roeger eliminate the income share of capital, which is why I have formulated equation (6) in terms of the income share of labour. Solow eliminates the income share of labour. He does not explain why he does that. Presumably, the reason is that he has no access to adequate data on the cost of labour. As we will see, Hall and Roeger eliminate the income share of capital for a similar reason. Solow complains about the bad quality of the data he uses for the income share of capital.[3] As theoretically speaking he could have eliminated that share instead of the share of labour, this suggests that the data of labour cost to which he had access were even worse or that he had no access to such data at all. By imposing the assumption of constant returns to scale, he could do without such data. This means that in effect the assumption of constant returns to scale functions as an empirical tractability assumption.

Again one can raise the question whether or not Solow believed the assumption to be true. This was probably not the case. Consider the following passage: ‘It proved possible to restrict the experiment to private non-farm economic activity. This is an advantage (a) because it skirts the problem of measuring government output and (b) because eliminating agriculture is at least a step in the direction of homogeneity [constant returns to scale]'. (Solow 1957: 314) This passage implies that Solow believes the assumption of constant returns to scale to be false with regard to agriculture. Furthermore, it suggests that he believes that the assumption may be false for other industries as well, but to a lesser degree. Solow could be read as imposing the assumption for reasons of applicability and negligibility. It is relatively clear that applicability considerations play a role here. Solow excludes agriculture, because he does not believe the theory including the assumption of constant returns to scale to be applicable to that industry. If negligibility were the reason for imposing the assumption, its effects would have to be (believed to be) negligible. This, however, is a rather optimistic interpretation. It is somewhat speculative to claim that 'to a lesser degree' means 'to a negligible degree'. Solow also claims that imposing the assumption of constant returns is 'practically unavoidable' (1957: 317). This fits well with the idea that the assumption was imposed for reasons of tractability. Overall, I conclude that the assumption is imposed for reasons of both applicability and empirical tractability.

Just as Solow, Hall and Roeger assumed constant returns to scale. As explained earlier, they used the assumption to eliminate the income share of capital rather than that of labor. Once more unobservability plays an important role. With respect to an equation that is similar to equation (1), Hall makes the following claim: '[T]his relation is not directly usable because the shadow value of capital, r [w_K], is not generally observed' (1988: 926). This does not mean that capital cost data cannot be constructed. Instead, it means that they are hard to construct and that once constructed they are frequently taken to be unreliable (Fisher 1987). By eliminating the income share of capital, this problem is (solved by being) circumvented. Note that Hall does not believe that returns are in fact constant. He finds high values of the mark-up in industries with low profits. He explains this in terms of a failure of the assumption of constant returns to scale (Hall 1987: 432).

Presumably, Roeger imposes the assumption of constant returns to scale for exactly the same reason as Hall. There is, however, one difference. In spite of the fact that Roeger eliminates the income share of capital from his estimation equation, he does need data on the level of capital cost. He creates them by combining Hall's data on the quantity of capital with his own calculations of the price of capital. The reason for this is that he needs to calculate changes in capital cost in order to calculate the difference between the primal and the dual Solow residual (see the definition of $\Delta x'$). The consequences of this, however, are less severe than they would be if data on the level of capital cost would be used directly as input of the estimation equation. The reason for this is that even if the data one constructs on the level of capital cost are unreliable, the changes that one calculates on the basis of these data may be reliable. As long as one uses a consistent method for calculating the level-data, this is very well possible. If the constructed level-data correlate well with the true values of the variable, relative changes calculated from the former will come close to those that would follow from the latter. Once more, we see that an assumption is imposed in order to circumvent problems related to the unreliability of data that threaten to make the measurement problem intractable. All in all, it is save to conclude that the assumption of constant returns to scale is an empirical tractability assumption for Solow, Hall, as well as for Roeger.

3.3 Constant productivity growth

The third assumption discussed in the previous section, that of constant productivity growth [A3: CPG], is used by Hall only. Once more unobservability provides the ulterior motive for imposing this assumption. Hall notes that if productivity growth would be known one could just calculate the mark-up from equation (10). He goes on to say that 'in practice the rate of productivity growth will not be known' (Hall 1988: 927). The phrase 'will not be known' is equivalent to 'is unobservable', as is clear from the context. Without this assumption, Hall's estimation equation (10) would contain two unknown variables, the mark-up and productivity growth. Given this assumption, productivity growth can just be equated to the constant of the regression and only one unknown remains. This means that the equation can be solved in virtue of the assumption. All this suggests that the assumption is an empirical tractability assumption, just as the two previous assumptions.

There are some considerations, however, that count against this interpretation. Hall writes for instance that his work 'proceeds on the assumption that the correlation [between productivity shits and the business cycle] is small enough to be ignored' (1986: 291). This suggests negligibility as a reason for imposing the assumption. On the other hand, the assumption of negligibility may be a false one, as might be implied by Hall's claim with regard to the alternative hypothesis that the correlation is substantial that 'the truth lies somewhere between the polar cases' (1986: 291). I regard the following passage as decisive for this issue:

> As the evidence now stands, one has a choice between these two very different views, both consistent with the principal evidence. Prior beliefs about the plausibility of large exogenous shocks in productivity are the primary basis upon which the choice has to be made. My own view is that productivity shocks … are not an important source of aggregate fluctuations'.
>
> (1986: 322)

Productivity shocks are unimportant as a source of aggregate fluctuation. In other words, they are negligible according to Hall. Those who support the real business cycle theory will disagree. For them the assumption will be an empirical tractability assumption, perhaps even a tractability assumption that conflicts with their views to such a large extent that they would be unwilling to use it. For Hall, however, the assumption is a negligibility assumption.[4]

As will be clear from the previous section, there are some complications concerning the relation between the dependent variables and the error term of the equation. Because of this, the instrument variable estimation method has to be used. This requires Hall to find an instrumental variable that has to meet certain conditions of its own. He recognizes that this is problematic, as will be clear from the following passage: 'It is a challenge to find instruments that are plainly exogenous on all views of macroeconomic fluctuations and that also have large enough influences on employment and output so that the test is powerful.' (Hall 1988: 932) Roeger criticizes Hall's method for the fact that it requires instrumental variable estimation. He qualifies the conditions that instruments have to meet as 'strong', claims that there are 'obvious difficulties' involved in finding suitable instruments, and portrays the fact that his method does not require instruments as an advantage (Roeger 1995: 328). Furthermore, he claims that '[p]oor instruments could be a main reason for a positive upward bias with Hall's method' (1988: 325). In addition to this he writes: 'I … shall argue that this approach is in some respects superior to the method suggested by Hall, since it does not require the use of instruments that are very hard to select' (1998: 318).

3.4 A constant mark-up

The assumption of a constant mark-up, [A4: CM], is imposed by both Hall and Roeger. In Hall's case, however, the imposition of the assumption is related to the identifying conditions of the instrumental variable. This means that its imposition is due to the assumption of constant productivity growth. The error term of Hall's estimation equation is equal to the random element of productivity growth. Because of this, the identifying conditions mentioned are as they are. If the mark-up would not be constant, it would most likely have a stochastic element. If it did, the error term of the estimation equation would be a composite of two stochastic elements. This in turn could mean that the identifying conditions do not hold anymore (the new element might introduce new correlation between the error term and an independent variable). Depending on the assumption made about the development of the mark-up, it would perhaps be feasible to formulate more complicated identifying assumptions. However, even if this were possible it would be of little help, as the additional complexity would make it even more difficult if not impossible to find suitable instruments.

It is difficult to determine whether or not Hall believes the mark-up to be constant in fact. On the one hand, he defends the assumption in ways that suggest he regards it as a negligibility assumption. Consider the following passages: 'Another assumption that I make is that a firm's mark-up ratio … can reasonably be approximated as a constant over time', and 'my work proceeds on the assumption that the correlation [between the business cycle and the mark-up] is small enough to be ignored' (Hall 1986: 290, 291). On the other hand, he emphasizes that, even though his estimation method requires the assumption of a constant mark-up, the test that he develops for perfect competition does not require this assumption (Hall 1988: 925, 927, 934). If Hall does not believe that the mark-up is in fact constant and if he cannot do anything but impose the assumption that it nevertheless is constant, the assumption is an empirical tractability assumption. I am inclined to believe the assumption is a tractability assumption rather than a negligibility assumption. What is striking in comparing this case with that of [A3: CPG] is that there Hall mentions his own view whereas here he uses a formulation that creates more distance between him and the assumption: rather than 'My own view is' he writes 'my work proceeds on the assumption that'. The latter is consistent with a qualification such as 'but my hunch is that the correlation is large enough to create a substantial bias'.

Roeger's case is different. He assumes that the mark-up is constant over time, even though he does not have to do so. There is, however, a tradeoff involved. Oliveira Martins *et al.* (1996) provide an application of Roeger's method without the assumption. The effect of relaxing the assumption is that the estimation equation becomes somewhat more complex. Thus its estimation requires more data.[5] Roeger defends the assumption first in terms of the goal he has: 'Since I want to demonstrate that even a simple variant of imperfect competition can help to reconcile price- and quantity-based productivity measures, I follow Hall and assume constant remarks' (Roeger 1995: 318). As negligibility is a matter that is relative to the purposes of the agent, this defence is consistent with the assumption being a negligibility assumption. Subsequently, he continues his defence in a footnote as follows: '[T]he evidence so far does not convincingly refute the assumption of acyclical mark-ups. … Given the weak empirical evidence in favor of pronounced cyclical mark-up fluctuations … my simplifying assumption seems not too strongly at odds with the data' (Roeger 1995: 318). The phrase 'not too strongly at odds with the data' together with the fact that rather than fluctuations *per se* Roeger argues against pronounced fluctuations confirm the impression that Roeger regards the assumption as a negligibility assumption.

The results of the investigation thus far are summarized in Table 1. The most immediate conclusion that can be drawn from them is that the amended Musgrave-Mäki typology is very useful for classifying the assumptions of the methods under consideration.

Solow, Hall and Roeger discuss a lot of other problems such as measurement problems and potential misspecifications. Many of the remarks they make in relation to those problems support the general thrust of this paper. They suggest that the assumptions they make fit the amended

Table 1 Types of assumptions

Assumption	*Unobservable*	*Economist*	*Reason for imposing the assumption*
A1: PC	Marginal cost	Solow	Empirical tractability
A2: CRS	Labour cost	Solow	Empirical tractability
	Price of capital	Hall	Empirical tractability
	Price of capital	Roeger	Empirical tractability
A3: CPG	Productivity growth	Hall	Negligibility (empirical tractability)
A4: CM	Productivity growth	Hall	Empirical tractability (negligibility)
	XXX	Roeger	Negligibility

(Less likely alternative interpretation)

Musgrave-Mäki typology. Solow assumes for instance that labour and capital always suffer unemployment to the same percentage and comments: 'This is undoubtedly wrong, but probably gets closer to the truth than making no correction at all' (1957: 314). Solow does not have data on capital utilization. He can only take this factor into account by making some simplifying assumption about it. This suggests that the assumption is an empirical tractability assumption. Solow also makes this remark: 'In general I limited myself to two-parameter families of curves, linear in the parameters (for computational convenience)' (1957: 318). Arguably, the assumptions about the form of the function involved are empirical tractability assumptions, as computational convenience is a pragmatic matter that arises – at least in this case – primarily when one tries to apply the model to reality.

As an exhaustive discussion of such remarks would be tedious, I mention only some remarks made by Roeger that are important for understanding the way in which he views the relation between his method and that of Hall. Having argued in favour of a simplifying assumption of his own (in terms of negligibility), Roeger attacks Hall for neglecting 'one important measurement problem' – compensated variations in work effort. Roeger himself is able to circumvent this problem because of the fact that '[m]easurement errors inflicted on both SR_t and SRP_t [the dual SR] will exactly cancel in this case' (cf. equation 10). He claims that '[t]he importance of this type of measurement problem is downplayed by Hall'. In addition to this, he argues that Hall's neglect of this issue results in an upward bias of his estimates. A final aspect of Roeger's discussion of this issue that is worth noting is that he puts it in terms of (un)observability. Hall cannot take variation in work effort into account because it is unobservable. Roeger does not need to take it into account, because the factor drops out of his equation (Roeger 1995: 322–3).

4 UNOBSERVABILITY, TRACTABILITY AND PROGRESS

Negligibility is a relatively pragmatic consideration for imposing a false assumption. It does not need to conflict with truth as the most important epistemic aim of the economist making the assumption. Even though truth is the main aim, it surely is not the aim at all costs. If truth in the sense of descriptive accuracy is not compromised too much, imposing a false assumption that has only negligible effects does not do much harm. Tractability is a more important consideration from this perspective. Tractability assumptions often pertain to factors the effects of which are not negligible. The fact that their imposition is usually unavoidable already suggests that they may have significantly distorting effects. This raises the question why economists impose tractability assumptions.

In the case under investigation, unobservability is the underlying reason. Solow assumes perfect competition due to the unobservability of marginal cost. Hall and Roeger assume constant returns to scale because of the unobservability of the price of capital. Hall assumes that productivity growth is constant because it is unobservable. The unobservability of productivity growth is also a reason for him to assume that the mark-up is constant. In the remainder of this section, I discuss how the issue of unobservability arises in empirical industrial organization and argue that many economists regard relaxing tractability assumptions as a way of achieving progress.

For an outsider it is often surprising that economists claim to deal with unobservable variables. Economists deal mostly with issues everybody encounters in daily life. Prices including interest rates and exchange rates, costs, inflation, firms, and the like are things most people will be familiar with. In light of this consideration, Mäki (1998) has claimed that economic entities are what he calls commonsensibles. Thus there is reason to ask what it means to an economist that something is unobservable. Let us consider marginal cost. Economists do not regard costs *per se* as unobservable. The costs of labour, for instance, are reported by businesses and collected by statistical agencies. Marginal cost, however, is a theorist's concept. It is the cost of the latest unit produced or used. The value of marginal cost cannot easily be computed from cost data that are reported. Economists often assume that the relation between marginal cost and the number of products produced is nonlinear. This means that marginal cost is not the same as average cost, except under special circumstances – notably perfect competition.

A somewhat different but related explanation of the unobservability of the price of capital can be given. Supposing for the moment that capital consists of machines, the unobservability of the price of capital makes intuitive sense. One may know the price of the machine. This means that in some sense the price of capital is observable. However, this is not the price of capital per product. It is difficult to obtain data on this, as the machine is used in the production of *many products* over *several years*. Both (italicized) aspects make it extremely difficult to decide what part of the cost of the machine should be assigned to a particular product. One of the problems involved is that it is often not known for how long exactly the machine will be used. Of course, accounting conventions exist concerning depreciation and related matters. These, however, do not always uniquely fix the price of capital. Furthermore, different firms may use different accounting conventions all of which may deviate from the ideal of economic theory in various ways. Problems are exacerbated by the fact that it is often hard even to determine total capital cost. Total revenue minus total labour and material cost is equal to the sum of capital cost and profit. As is well known, firms can manipulate their profit figures to a fair extent in spite of existing regulations. All this means that the issue of measuring the price of capital is a matter of staggering complexity, and I have not even mentioned costs of maintenance of machines, the fact that the value of a machine may decrease due to use rather than or in addition to the passage of time, as well as the many other sorts of capital goods that can be distinguished.

The unobservability of productivity growth is perhaps most intuitive. Although one may be aware of the fact that productivity has increased over time, such an increase will not be apparent from figures used on invoices and in (basic) accounting procedures. This means that in contrast to the previous two cases, it may happen that there are no data on productivity growth at all. As will be clear by now, unobservability in economics does not concern the question whether there is really something particular there to be measured in the first place as it does in physics. Instead, unobservability in economics is largely a matter of *measurability*, or more specifically with the question of *whether the right data are used or supplied by economic agents themselves*. The question at issue is whether the values of the properties on which data are needed *can be reliably measured as they are conceived of by economic theory*. Reliability is a problem when different economic agents use different methods for measuring the value of the property at hand. The qualification ‘as they are conceived of by economic theory’ is relevant when the common sense or even the accountants’ conception of a property such as cost, deviates from the theorist’s conception (for example of cost as marginal cost). Thus unobservability in economics is at least to some extent an artefact of economic theory.[6]

As we have seen in the previous section, economists talk of the values of variables they regard as unobservable as unknown. In addition to that, they sometimes claim that they are not directly observable. What they mean by this is not that the property is not directly observed *per se*, but instead that *the true value of the property as it is conceived of by economic theory is not directly observed*. This usage fits well with the solutions that are put forward for measuring productivity growth and the mark-up. The economists discussed try to infer the values of these variables from the values of other values by postulating relations that are believed to hold between them. More specifically, they start from a model that represents the relations between inputs and output in mathematical terms. The model is called a production function. Their aim is to measure the value of a variable that is unobservable. Supposing for the moment that all other variables of the model are observable, this is a straightforward matter. Once the required data have been collected, the model can compute the value of the unobservable variable from these data.

A complication arises because not all assumptions are (believed to be) true. Rather than concluding that the value of the targeted value cannot be computed at all, however, Solow, Hall and Roger find ways around the problem of unobservability. They solve what are in more general terms tractability problems by imposing false assumptions that involve simplifications of reality. Such assumptions are justified in terms of second-order assumptions, in this case tractability assumptions. Economists such as the ones under consideration are aware that the simplifications can have distorting effects on the outcomes of their models. In light of this, they perceive the relaxation of the assumptions justified in terms of tractability assumptions as an important way of achieving progress.

One wonders whether there is not a more direct way of assessing the quality of one's estimates than checking the truth of the assumptions. Roeger does have some independent means of checking the validity of his estimates and comparing it with the validity of Hall's estimates. He uses the same data as Hall (1988), which ensures that the results can be compared. Roeger's estimates suggest substantially lower mark-ups. Roeger argues first that his estimates are better than those of Hall because the statistical significance of his estimates is much higher than that of Hall's. Secondly, he argues in favor of his own estimates by comparing them to results obtained by other methods that are beulieved to be reliable. Roeger writes: 'My results are also more in line with cross section studies such as, for example, the study by Bresnahan (1981) on the US automobile industry'. (1995: 325)

It is impossible, however, to compare the estimates directly to the true values and the two relatively direct ways of assessing their quality are not sufficient. This makes the truth of the assumptions all the more important. In effect, the economists of the case under investigation maximize the accuracy of the measurements by maximizing the truth of the assumptions of their measurement methods.[7] This can be gathered from the other considerations Roeger puts forward in favour of his method over and against that of Hall. Roeger's main argument is that his method does not require instrumental variable estimation. Related to this, he does not need to assume that productivity growth is constant over time. By invoking the dual Solow residual and subtracting it from the primal SR Roeger is able to solve an empirical tractability problem that may have had distorting effects on Hall's estimates. Roeger's main argument appears in several guises. First, he claims that the identifying conditions that the instruments have to meet are strong. Second, he claims that is difficult to find suitable instruments. Third, he claims that Hall's estimates may be biased because his instruments do not meet the identifying conditions in full. Fourth, he claims that his own method is superior to that of Hall in virtue of the fact that his method does not require instrumental variables that have the disadvantages mentioned.

In terms of the ideas that I have defended thus far, all this boils down to the thesis that Roeger's measurement method is better than that of Hall because it requires fewer tractability assumptions. This fits well with the though that (some) economists aim at building models that are accurate representations of the real world in the sense that all representational aspects be correct (in contrast to the sense that they represent all aspects correctly). This realist attitude comes out nicely in some of the things these economists claim themselves. Roeger sees the situation in which all assumptions are fulfilled or true as 'ideal' (1995: 317).[8] From his paper, it seems warranted to infer that when the situation is not ideal, one should aim at making it ideal by relaxing (an) assumption(s) to make the set of assumptions more realistic. In other words, it can be made 'ideal' by relaxing 'idealizing assumptions'. Solow makes the following remark: 'Naturally, every additional bit of information has its price' (1957: 312). This sums up nicely the core message of my paper. Measuring an unobservable variable requires data and a representational system. Although some aspects of the representation must be true, the price of measuring an unobservable variable consists in imposing some unrealistic assumptions.

To what extent does the view defended here generalize? I suggest that it holds at least for most of applied industrial organization. Support for this claim comes from one of the leading economists in that field. Sutton (2000, 2002) uses the terms 'unobservability' and 'tractability'. He claims that unobservables pose a serious problem for economic research. He describes the general strategy for addressing the related tractability problems as follows: '[W]e aim to handle the unobservables by designing the theory in such a way as to allow us to work round them' (2002: 57). What is missing in this characterization is that apart from learning to live with tractability problems and addressing them by imposing tractability assumptions one can try to solve the tractability problems, i.e. one can attempt to find a way of relaxing the related tractability assumptions. I am sure, however, that he would agree. Thus, capturing progress in terms of tractability provides a nice summary statement of the ways in which economists themselves assess their models.

Frank A. Hindriks
Erasmus University Rotterdam
hindriks@fwb.eur.nl

ACKNOWLEDGEMENTS

Earlier versions of this paper have been presented at the summer school *Economics at the Turn of the Millennium: Pondering Progress* (July 1999; San

Sebastian, Spain), at the *Arctic Workshop on Theorizing and Experimentation in Economics* (December 1999; Rovaniemi, Finland) and at the *2004 INEM conference* (August 2004; Amsterdam, the Netherlands). I thank the participants in these events as well as Uskali Mäki and Roberto Lazzari for their helpful comments.

NOTES

1 According to duality theory, profit maximization has the same implications in terms of prices and quantities of inputs and outputs as cost minimization. The primal SR is derived from a production function assuming profit maximization. The dual SR is derived from a cost function assuming cost minimization. Because of the equivalence of profit maximization and cost minimization – under certain assumptions (see Takayama 1985: 146–7) – duality theory can be used to show that the primal and dual SR are equivalent.

2 Even though Hall and Roeger do not assume perfect competition on the market for outputs they do assume perfect competition on the market for inputs, just as Solow does. It is quite plausible that this assumption is a tractability assumption for them as well.

3 Solow makes the following remarks about the quantity of capital and the income share of capital respectively: 'The capital time series is the one that will really drive the purist mad' and 'The share-of-capital series is another hodge-podge, pieced together from various sources and ad hoc assumptions' (1957: 314; see also note 1 on p. 312).

4 Hall is, however, well aware of the fact that apart from the first-order assumption that productivity growth is constant the second-order assumption that deviations from constancy have a negligible effect can be false as well. In another paper of his, he presents some evidence that suggests that this is indeed the case, although he remarks that 'the evidence is far from definitive' (Hall 1987: 443; see also p. 440).

5 The new variable should be an indicator of the business cycle. Data for such a variable are readily available. In fact, Roeger estimates a variant of his own equation that involves such a variable. However, he adds it in an *ad hoc* manner and does not derive the new estimation equation from a more general one as Oliveira Martins *et al.* (1996) do.

6 The weakest sense in which the term 'unobservable' is used is when the value of the property in question is not measured in fact. Roeger seems to use the term in this sense when he discusses variation in work effort. It can be questioned whether this use of the term is still adequate. Observability is a modal notion, i.e. it pertains to the possibility of observing something. In the context of economics, it can be connected to another modal notion, measurability, as I do in the main text. When it is used in relation to actual measurement rather than possible measurement, however, the modal aspect of the notion is lost altogether.

7 Zamora Bonilla (1999) has argued that increasing the plausibility of assumptions is an especially suitable strategy for economists as compared to, for instance, physicists, due to the problems economists face in empirical research.

8 This observation is important because it runs counter to a strong current in economics inspired by Milton Friedman (1953) that denounces increasing the realisticness of assumptions as a relevant pursuit. When Friedman's paper was

published, many economists considered the realisticness of assumptions to be important. De Marchi writes for instance the following about a seminar at the London school of economics in the fifties: 'Models were examined for the realism of their assumptions and for internal consistency' (De Marchi 1988: 143). As another example, Martin Bronfenbrenner makes the following remark about the 'standard textbook position': '[N]o theory can be "useful" or command confidence if its underlying assumptions are "unrealistic"' (1966: 14). The current investigation makes clear that the realisticness of assumptions is still an important consideration in model-building today, at least in parts of economics.

REFERENCES

Brensnahan, T.F. (1989) 'Empirical analysis of industries with market power', in R. Schmalensee and R. Willig (eds) *Handbook of Industrial Organization*, Vol. II, North-Holland, pp. 1012–57.

Bronfenbrenner, M. (1966) 'A 'middlebrow' introduction to economic methodology', in S.R. Krupp (ed.) *The Structure of Economic Science: Essays on Methodology*, London: Prentice-Hall, pp. 5–24.

Church, J.R. and Ware, R. (2000) *Industrial Organization: A Strategic Approach*, Boston: McGraw-Hill.

De Marchi, N. (1988) 'Popper and the LSE economists', in N. De Marchi (ed.) *The Popperian Legacy in Economics*, Cambridge: Cambridge University Press, pp. 139–66.

Fisher, F.M. (1987) 'On the misuse of the profit-sales ratio to infer monopoly power', *RAND Journal of Economics* 18: 384–96.

Friedman, M. (1953) 'The methodology of positive economics', in *Essays in Positive Economics*, Chicago: Chicago University Press, pp. 3–43.

Hall, R.E. (1986) 'Market structure and macro-economic fluctuations', *Brookings Papers on Economic Activity* 2: 285–338.

Hall, R.E. (1987) 'Productivity and the business cycle', *Carnegie-Rochester Conference Series on Public Policy* 27: 421–44.

Hall, R.E. (1988) 'The relation between price and marginal cost in US industry', *Journal of Political Economy* 96: 921–47.

Hindriks, F.A. (2005) 'Tractability assumptions and the Musgrave-Maki typology', *Mimeo*.

Hulten, C.R. (1986) 'Productivity change, capacity utilization, and the sources of efficiency growth', *Journal of Econometrics* 33: 31–50.

Mäki, U. (1998) 'Aspects of realism about economics', *Theoria* 13: 301–19.

Mäki, U. (2000) 'Kinds of assumptions and their truth. Shaking an untwisted F-twist', *Kyklos* 53: 303–22.

Mayer, T. (1999) 'The domain of hypotheses and the realism of assumptions', *Journal of Economic Methodology* 6: 319–30.

Musgrave, A. (1981) '"Unrealistic assumptions" in economic theory: the F-twist untwisted', *Kyklos* 34: 377–87.

Oliveira Martins, J., Scarpetta, S. and Pilat, D. (1996) 'Mark-up-pricing, market structure and the business cycle', *OECD Economic Studies* 27: 71–105.

Roeger, W. (1995) 'Can imperfect competition explain the difference between primal and dual productivity measures? Estimates for US manufacturing', *Journal of Political Economy* 103: 316–30.

Solow, R.M. (1957) 'Technical change and the aggregate production function', *Review of Economics and Statistics* 39: 312–20.

Sutton, J. (2000) *Marshall's Tendencies. What Can Economists Know?*, Cambridge (MA): MIT Press.
Sutton, J. (2002) 'Marshall's tendencies: a reply', *Economics and Philosophy* 18: 55–62.
Takayama, A. (1985) *Mathematical Economics*. 2nd edn, Cambridge: Cambridge University Press.
Zamora Bonilla, J.P. (1999) 'Verisimilitude and the scientific strategy of economic theory', *Journal of Economic Methodology* 6: 331–49.

APPENDIX

As explicated in the main text – see equation (6) – the traditional formula for calculating the SR is:

$$SR = \Delta q - \Delta k - \alpha_N^P(\Delta n - \Delta k) \tag{i}$$

From equation (9), it follows that, if competition is not perfect, the SR equals:

$$SR = (\mu - 1)\alpha_N^P(\Delta n - \Delta k) + \theta \tag{ii}$$

When equations (i) and (ii) are set equal to each other an expression for $\alpha_N^P \Delta n$ can be derived:

$$\alpha_N^P \Delta n = \frac{1}{\mu}(\Delta q - \theta) + \left(\alpha_N^P - \frac{1}{\mu}\right)\Delta k \tag{iii}$$

If this is inserted in the Solow residual, (i), the following results:

$$SR = \left(1 - \frac{1}{\mu}\right)(\Delta q - \Delta k) + \frac{1}{\mu}\theta \tag{iv}$$

This can be rewritten as:

$$SR = L(\Delta q - \Delta k) + (1 - L)\theta \tag{v}$$

The dual or price-based Solow residual, SRP, follows from a largely similar derivation that starts from a cost function instead of a production function, and is given by:

$$SRP = -L(\Delta p - \Delta w_K) + (1 - L)\theta \tag{vi}$$

As can be seen from (v) and (vi), both the primal and the dual Solow residual consist of a mark-up and a technology component. In a situation of perfect competition these two measures should be equal to each other. When the two differ this can be explained by the presence of a mark-up. This can be seen when the dual measure is subtracted from the primal measure. The technology component drops out of the equation and the mark-up component remains:

$$SR - SRP = L(\Delta q - \Delta k) - \{-L(\Delta p - \Delta w_K)\} = L\Delta x \tag{vii}$$

where Δx is defined as:

$$\Delta x \equiv (\Delta q - \Delta k) + (\Delta p - \Delta w_K) \tag{viii}$$

The left-hand side, the difference between the two Solow residuals, is designated as Δv, and equals:

$$\Delta v \equiv (\Delta q + \Delta p) - \alpha_N^P(\Delta n + \Delta w_N) - \left(1 - \alpha_N^P\right)(\Delta k + \Delta w_K) \tag{ix}$$

Using these notations, the equation estimated by Roeger (1995) follows:

$$\Delta v = L\Delta x + \omega \tag{x}$$