Available online at www.sciencedirect.com
ScienceDirect
Journal of Macroeconomics 30 (2008) 671–686
Journal of MACROECONOMICS
www.elsevier.com/locate/jmacro

σ: The long and short of it ☆

Robert S. Chirinko *

Department of Finance, 2421 University Hall, University of Illinois at Chicago, 601 South Morgan (MC 168), Chicago, IL 60607-7121, United States

Received 21 April 2007; accepted 20 October 2007
Available online 21 November 2007

Abstract

Research on the elasticity of substitution between capital and labor – σ – has been proceeding for 75 years. While there is clearly a strong case for the importance of σ in the analysis of growth and other economic issues, much less agreement exists on the value of σ. This paper offers some perspectives on prior estimates of σ, emphasizing the fundamental tension between the short-run data that are available and the long-run parameter that is required. Estimates of σ based on various short-run and long-run models are discussed and, while the estimates range widely, the weight of the evidence suggests a value of σ in the range of 0.40–0.60. There is little evidence to sustain the assumption of a Cobb–Douglas production function.

JEL classification: E22; C20

Keywords: Production; Function; Capital formation; Economic growth; Estimating long-run relations

1. Introduction

Research on the elasticity of substitution between capital and labor – σ – has been proceeding for 75 years. Over this extended period, the importance of σ in economic growth has been demonstrated with respect to, among other issues, the possibility of perpetual

☆ This title represents not only one of the major themes of this paper, but also an American idiom meaning "the bottom line" or "the crucial point."

* Tel.: +1 312 355 1262; fax: +1 312 413 7948.
E-mail address: chirinko@uic.edu

0164-0704/$ - see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmacro.2007.10.010

growth or decline,[1] the level of per capita income (Klump and de La Grandville, 2000), growth in per capita income (de La Grandville, 1989), the rate of return on capital (King and Rebelo, 1993; Mankiw, 1995), the speed of convergence (Klump and Saam, 2007; Turnovsky, 2008), and the relative roles of productive factors and technical efficiency in explaining differences in per capita income (Caselli, 2005). The theoretically-oriented papers presented at this conference offer fresh insights into σ's central role in growth models.

Apart from long-run growth, σ looms large in analyzing short-run fluctuations and tax policies. Monetary policy affects real activity by altering interest rates, and σ influences the quantitative importance of this part of the monetary transmission mechanism. Implicit assumptions about σ are important in dynamic stochastic general equilibrium models that examine short-run fluctuations. Most of these models are based on a Cobb–Douglas production function with its σ equal to one (e.g., Farmer, 1997); this assumption may bias upward the response of real variables to interest rates and other prices if the true value of σ is less than one. The simulated effects of fiscal policies are sensitive to the assumed value of σ. For example, Engen et al. (1997, Tables 2A and 5), show that, due to a shift from an income tax to a consumption tax, increases in steady-state net output are 3.8, 6.8, or 9.5 percent depending on whether σ is 0.5, 1.0, or 1.5, respectively.[2]

While there is clearly a strong case for the importance of σ in the analysis of growth and other economic issues, much less agreement exists on the value of σ. This paper offers some perspectives on estimation strategies for σ, tabulates key results that have appeared in the literature, and attempts to develop a consensus range of values for σ.

The paper begins in Section 2 with some background on the history of σ and the CES production function. The remaining sections touch on some of the pitfalls and obstacles that face the applied econometrician when estimating σ and on the fundamental tension between the short-run data that are available and the long-run parameter that is required. Section 3 explores the role of σ in short-run models of capital accumulation and the difficulties encountered by both explicit models (Brainard–Tobin's Q, Euler equation, and direct forecasting) and implicit models in generating reliable estimates. A similar task is undertaken in Section 4 for long-run models (unadjusted first-order condition, cointegration, and interval-difference) that directly estimate the first-order condition for capital with appropriate adjustments or assumptions to account for the short-run nature of the data. A partial reconciliation is achieved among the disparate estimates from the three models. Section 5 concludes with some observations on estimates of σ and lingering issues.

Before turning to these issues of theory and empirics, a bit of background will prove useful.

2. Some background

The concept of the elasticity of substitution between capital and labor was introduced nearly 75 years ago independently by John Hicks and Joan Robinson. Hicks can claim

[1] The possibility of perpetual growth was first noted by Solow (1956), Pitchford (1960), and Akerlof and Nordhaus (1967) and has featured in the more recent analyses of Barro and Sala-i-Martin (1995), de La Grandville (1989), Klump and Preissler (2000), Klump and de La Grandville (2000), and de La Grandville and Solow (2004).

[2] Additional tax simulation models are analyzed in Chirinko (2002, Table 2).

priority and came to the concept while studying wages and the allocation of national income between the factors of production. His crucial insight was that the impact of the capital/labor ratio on the distribution of income (given output) could be completely characterized by the curvature of the isoquant.[3] For a neoclassical production function relating output (Y) to capital (K) and labor (L), $Y = H[K,L]$, he formalized the elasticity concept in terms of the following partial derivatives,

$$\sigma \equiv (H_K H_L)/(H_{KL} Y). \tag{1}$$

Robinson (1933/1959, p. 256) introduced an alternative formulation of the substitution elasticity,

$$\sigma \equiv -((\mathrm{d}K/K) - (\mathrm{d}L/L))/((\mathrm{d}P^k/P^k) - (\mathrm{d}P^L/P^L)), \tag{2a}$$

$$= -(\mathrm{d}\log(K/L))/(\mathrm{d}\log(P^k/P^L)), \tag{2b}$$

where P^k and P^L are the factor prices for capital and labor, respectively. The relations between Eqs. (1) and (2) are not immediately obvious, and Hicks demonstrated in the second edition of his book of wages (1963, section III, "notes on the elasticity of substitution", sub-section 1) that the two formulations are equivalent (under constant returns to scale). While Hicks can claim priority, Robinson has captured the attention of posterity, and her definition has proven the more popular and durable.

The mention of Joan Robinson in a paper on production functions naturally stirs memories of the Cambridge Capital Theory Controversies. While this debate touches on a multitude of issues, the ones discussed in Schefold (2008) most relevant for present purposes are the estimation of production function parameters and the measurement of the capital stock. The estimation issue has been raised by Shaikh (1974) and Felipe and Holz (2001), who argue that the accounting identities inherent in value added data necessarily lead to a Cobb–Douglas production function. This argument is called into question by those studies with aggregate data that report estimates of σ much less than one, and it has been forcefully challenged by Solow (1974, 1987) on empirical and theoretical grounds. Regarding the latter point, Solow (1987, p. 20), shows that, when factor shares are constant, any production function can be represented as the product of a Cobb–Douglas production function and the production function suitably normalized. The second issue concerning capital stock measurement strikes this author as of small moment because the measurement issue turns on the rate of interest being sensitive to the capital stock through the marginal product of capital. However, in many general equilibrium models, the interest rate is determined by preference parameters invariant to movements in the capital stock and, in these models, equality between the rate of interest and the marginal product of capital is achieved by variations in non-capital factors of production. Moreover, in models of the open economy or of a panel of firms or industries, the interest rate is again independent of the capital stock. These and other aspects of the Cambridge Controversies seem to turn-on the consideration of aggregate quantities and aggregate production functions and, regardless of the arguments presented above, it remains unclear whether the Cambridge Controversies apply to data and parameters estimated from panel data.

[3] See Blackorby and Russell (1989, p. 882). These authors, as well as de La Grandville (1997), show that the substitution elasticity is no longer measured by the curvature of the isoquant when there are more than two factors of production.

As noted in the title of this conference, the second major analytic innovation occurred with the introduction of the constant elasticity of substitution (CES) production function by Arrow et al. (1961),

$$Y_t^* = F[K_t^*, L_t^* : A_t, B_t^K, B_t^L, \phi, \sigma], \tag{3a}$$

$$= A_t\{\phi(B_t^K K_t^*)^{[(\sigma-1)/\sigma]} + (1-\phi)(B_t^L L_t^*)^{[(\sigma-1)/\sigma]}\}^{[\sigma/(\sigma-1)]}, \tag{3b}$$

where Y_t^* is long-run real output that depends on (a) two choice variables – K_t^*, the long-run real capital stock (dated at the beginning of the period), and L_t^*, the long-run level of labor input; (b) three variables representing exogenous technical progress – A_t, neutral technical progress, and B_t^K and B_t^L, capital-biased and labor-biased technical progress, respectively; and (c) two parameters characterizing the production technology – ϕ, the capital distribution parameter, and σ. The superscript * reminds us that quantities entering the production function are long-run variables; that is, variables that have been adjusted to their optimal levels without incurring costly frictions.[4] Assuming that the firm maximizes profits and faces fixed input and output prices, we obtain the following first-order conditions for capital and labor, respectively, stated in log-linear form,

$$ky_{\mathrm{t}}^* = \gamma^{ky} - \sigma \mathrm{ucc}_t^* + v_t^{ky}, \tag{4a}$$

$$\ell y_t^* = \gamma^{\ell y} - \sigma \mathrm{uc}\ell_t^* + v_t^{\ell y}, \tag{4b}$$

where $ky_t^* \equiv \ln((K/Y)_t^*), \gamma^{ky} \equiv \sigma\ln(\phi), \sigma$ is the substitution elasticity, $\mathrm{ucc}_t^* \equiv \ln((P^k/P^Y)_t^*)$, (the Jorgensonian user cost of capital), $v_t^{ky} \equiv \ln(A_t^{[\sigma-1]}B_t^{K[\sigma-1]}), \ell y_t^* \equiv \ln((L/Y)_t^*), \gamma^{\ell y} \equiv \sigma\ln(1-\phi), \mathrm{uc}\ell_t^* \equiv \ln((P^L/P^Y)_t^*)$ (the user cost of labor or the real wage rate), and $v_t^{\ell y} \equiv \ln(A_t^{[\sigma-1]}B_t^{L[\sigma-1]})$.

These first-order conditions invite at least four comments. First, the only source of error is due to the technology shocks, and it is the interaction between a value of σ differing from unity and the technology shock that permits a stochastic element to enter the first-order conditions. When σ equals unity, this stochastic element disappears, and the first-order condition holds exactly. In this unrealistic case, other sources of error – for example, mismeasured variables – are needed to preserve a stochastic specification. Second, neutral and biased technical progress enter the first-order conditions symmetrically and thus are not separately identifiable. A similar identification problem exists with production functions, as demonstrated by Lau (1980). Moreover, it can be difficult to distinguish between the effects of biased technical progress and σ (Diamond et al., 1978). These problems force researchers to either assume away one type of technical progress or, as in Klump et al. (2007, 2008), to assume functional forms that permit identification. Third, the distinction between short-run and long-run elasticities is not easily accommodated in a production function. To allow for such a distinction, the production function would need to be embedded in a model of costly adjustment. The production function and its parameters described in Eqs. (3) and (4) describe long-run situations after adjustments have taken place. Fourth, the CES production function has been parameterized in different ways. For example, Klump and de La Grandville (2000) note that the formulation in Eq. (3) implicitly assumes a baseline point. They normalize the CES production function at this baseline point (by scaling the output and input variables to form consistent index numbers

[4] In his survey, Caballero (1999) refers to these long-run values as "frictionless."

and by an appropriate choice of factor income shares) such that CES functions with different σ's but otherwise the same parameters are tangent. This normalized CES function has been successfully estimated by Klump et al. (2007, 2008).

The production function (Eq. (3)) and associated first-order conditions (Eq. (4)) have been the primary vehicle for studying factor demands, and they have been the basis of most prior estimates. Thus, the traditional formulation will be used in the remainder of this paper. Moreover, based on the comparative advantage of the author, this study will focus on the first-order condition for capital.[5]

3. σ: The short of it

The fundamental tension – the available short-run data vs. the required long-run parameter – forces researches to take one of two paths in obtaining estimates of σ. One path leads to a focus on long-run relations with appropriate adjustment to account for the short-run nature of the data. This path is pursued in Section 4. The current section considers the alternative path that takes the short-run data as given and modifies the estimating equation to account for frictions and dynamics. All models consider capital accumulation, though they are equally applicable to the demand for labor as well.

3.1. Explicit models

Two approaches to modeling short-run data are considered here.[6] The first is based on an explicit optimization problem. These models require some friction to impede adjustment to the long-run steady-state, and the most popular assumption has been convex adjustment costs for accumulating capital. Such a modeling strategy, introduced by Eisner and Strotz (1963), has been in the literature nearly as long as the CES production function. Adjustment costs represent either external costs due to an upward sloping supply curve for capital goods or internal costs. The latter represent lost output from disruptions to the production process (as new capital goods are "broken-in" and workers retrained), additional labor for "bolting-down" new capital, or a wedge between the quantities of purchased and installed capital. These costs increase at an increasing rate, an assumption that plays a crucial role in explicit models. With linear or concave adjustment costs, the firm would pursue an all-or-nothing investment policy. Convexity forces the firm to think seriously about the future, as too rapid accumulation of capital will prove excessively costly. Alternatively, too little accumulation results in foregone profits.

Specifying adjustment costs as a quadratic function of investment with parameter $\zeta - (\zeta/2)(I_t^2/K_t)$, where I_t and K_t are investment and capital, respectively – and inserting this convex adjustment cost function into a dynamic optimization problem, we obtain the

[5] See Hamermesh (1993) for an extensive discussion of σ in the first-order condition for labor. Duffy and Papageorgiou (2000) and Mallick (2006) present recent estimates of σ directly from the CES production function; the former and latter studies report estimated σ's that tend to be greater than and less than unity, respectively. The working paper version of Klump et al. (2007) reports additional results from existing studies estimating a variety of models for the United States (Table 1) and non-US countries (Table 2).

[6] See Chirinko (1993a,b, 2002) for a fuller discussion of the explicit and implicit models and the empirical results discussed in this section.

following relation between investment costs today and production benefits today and tomorrow,

$$(I_t/K_t) = \delta + (1/\zeta)(E_t\{\Lambda_t\} - (P_t^I/P_t^Y)) + V_t^I, \tag{6a}$$

$$\Lambda_t \equiv \sum_{s=0}^{\infty} \rho^s \lambda_{t+s}, \tag{6b}$$

where (P_t^I/P_t^Y) is the price of new investment goods relative to the price of output, δ is the geometric rate of capital depreciation, and V_t^I is a stochastic error term representing shocks to the adjustment cost technology. The current and future benefits of a piece of capital acquired today are measured by the shadow price of capital, Λ_t, which depends on a constant discount rate, ρ, and, with a modification to account for adjustment costs, the marginal product of capital, λ_t. In explicit investment models, the σ parameter enters through λ_t.

Eq. (6a) is the dynamic equivalent of the simple decision rule for the optimal capital stock in Eq. (4a) in which the expected marginal benefits are equated to the costs of investing in period t. The marginal benefit is measured by the shadow price of capital, Λ_t. Owing to capital's durability, this is the discounted sum of the "spot" marginal revenue products (λ_{t+s}'s) over the life of the capital good as evaluated with information available in period t. The marginal costs are the sum of purchase costs and the sunk adjustment costs associated with investing. Since the sunk costs can not be recovered, they force the firm to look ahead when investing. Whenever there is a discrepancy between $E_t\{\Lambda_t\}$ and (P_t^I/P_t^Y), the firm has an incentive to change its capital stock, but its actions are tempered by the convex adjustment cost technology. The steeper is the adjustment cost function, the larger is ζ, and the more slowly investment responds.

The explicit model is appealing because it is internally consistent and recognizes the separate influences of expectations and dynamics, thus being immune from the Lucas Critique. For empirical researchers, the critical problem with developing an estimable equation is relating the unobservable Λ_t to observable variables.

There are three general solutions to this unobservability problem. One solution (the Brainard–Tobin Q model) equates $E_t\{\Lambda_t\}$ to data in financial markets. When the value of the firm as evaluated on financial markets exceeds the replacement cost of its assets, the first term on the right-side of Eq. (6a) is positive, and the firm acquires capital. A second solution (the Euler equation model) transforms $E_t\{\Lambda_t\}$ in an appropriate way so as to eliminate all but one of the future variables appearing in Eq. (6b). The resulting equation contains investment and relative prices in periods t and $t+1$, as well as λ_t, which depends on σ. The third solution (direct forecasting) specifies a VAR that includes λ_t and other variables related to λ_t (perhaps including the discount factor entering Eq. (6b) if it is stochastic), estimates the VAR parameters, projects a path for λ_{t+s}, $s \geqslant 1$, and then computes Λ_t. Based on information available at time t, this forecasted path is used to construct $E_t\{\Lambda_t\}$.

There are two important shortcomings of explicit models. First, the models have not usually been used to estimate σ. Indeed, σ is not recoverable from the Brainard–Tobin's Q model because the financial market data is a sufficient statistic for all information relevant for the capital formation decision. Price elasticities have been estimated with Euler equation models (e.g., Pindyck and Rotemberg, 1983a,b). Since their model is based on a translog technology, σ can not be inferred from their parameter estimates, but it should be noted that the price elasticity of capital estimates differ dramatically in the two studies from −2.93 to −0.13 in the 1983a and 1983b studies, respectively. In their direct forecasting model, Abel and Blanchard (1986) estimate models with different pre-set values of σ and find that the variability of the estimated $E_t\{\Lambda_t\}$'s is sensitive to σ. Second, the tight structure of these three explicit models is desirable insofar as the frictions incorporated into the optimization problem have been modelled properly. However, in addition to quadratic adjustment costs, capital accumulation may be affected by a variety of frictions – asymmetric and non-convex adjustment costs for capital, adjustment costs for investment, costly reversibility, finance constraints, and time-to-build and gestation lags. Unfortunately, a large body of empirical evidence calls into question the empirical performance of explicit models, thus casting suspicion on the underlying specification of frictions.

3.2. Implicit models

The alternative strategy is to estimate σ from models motivated by theory but whose specifications do not follow explicitly from an optimization problem. While specifications of implicit models vary widely, they generally relate investment spending to a price variable (frequently the Jorgensonian user cost of capital) and a quantity variable such as sales or output. The impact of frictions in temporarily impeding capital accumulation are represented by distributed lags in the user cost of capital (UCC_t) and sales variables (S_t) in the following specification:

$$(I_t/K_t) = \delta - \sum_{j=1}^{J} \alpha_j^{\mathrm{UCC}} \Delta \mathrm{UCC}_{t-j}/\mathrm{UCC}_{t-j} + \sum_{j=1}^{J} \alpha_j^{S} \Delta S_{t-j}/S_{t-j} + V_t^{I/K}, \tag{7a}$$

$$\sigma = \sum_{j=1}^{J} \alpha_j^{\mathrm{UCC}}, \tag{7b}$$

where $V_t^{I/K}$ is a stochastic error term. The derivation of this implicit investment equation begins with the first-order condition for capital (Eq. (4a)). We then assume that new investment is a distributed lag of changes in the optimal capital stock and that replacement investment is proportional to the existing capital stock. The price elasticity of capital equals the summation of the α^{UCC}'s in Eq. (7b) and, under a CES production function holding output constant, this price elasticity equals σ. The model described in Eq. (7) can also be derived by assuming that the capital stock is generated by the following distributed lag model on user costs and sales,

$$k_t = -\sum_{j=1}^{J} \beta_j^{\mathrm{UCC}} \mathrm{ucc}_{t-j} + \sum_{j=1}^{J} \beta_j^{S} s_{t-j} + v_t^{k}, \tag{8}$$

where k_t, ucc_t, and s_t represent the logarithms of the capital stock, the user cost of capital, and sales, respectively, and v_t^k is a stochastic error term. Eq. (7a) is obtained by first-differencing Eq. (8) and noting that $I_t/K_{t-1} = \Delta k_t + \delta$, where δ is the geometric rate of capital depreciation.

The table contains some key results from implicit models. The early and prominent studies by Jorgenson (1963) and Hall and Jorgenson (1967, 1971) were based on a Cobb–Douglas production function, and hence σ equals 1.00 by assumption. Eisner and Nadiri (1968) estimated σ freely, and found that the responsiveness of capital to its user cost was between 0.16 and 0.33. The gap has not been closed by subsequent research using aggregate investment data. A summary by Chirinko (1993a,b) found the elasticity estimates varied widely, but they tended to be small and less than 0.30. Recent work supports the latter finding for aggregate investment. Tevlin and Whelan (2003) find an elasticity of 0.18 for total aggregate investment, though a much larger value of 1.59 for investment in computers. These results are confirmed with UK data by Bakhshi et al. (2003), who report estimates of 0.32 and 1.33 for aggregate and computer investment, respectively. Also using UK aggregate data, Ellis and Price (2004) find that σ equals 0.44.

Aggregate data have several drawbacks. There is a limited amount of variation relative to industry or firm-level datasets, and thus parameters may be imprecisely estimated. Furthermore, problems of simultaneity, capital market frictions, or other sources of heterogeneity may bias the estimated elasticities. Simultaneity bias arises because a positive shock to investment demand will raise interest rates (embedded in the user cost) either because the supply of saving is upward sloping or the monetary authorities attempt to moderate fluctuations. In either case, the regression error term and user cost variable will be positively correlated, and the estimated elasticity biased toward zero. Capital market frictions affecting different classes of firms (e.g., those that are highly leveraged) and their sensitivity to price incentives have been documented in many studies (see Fazzari et al. (1988) and the survey by Hubbard, 1998).

To address these concerns, recent research has explored the price sensitivity of capital with large panel datasets (in some cases containing approximately 25,000 firm/year observations). Cummins and Hassett (1992) and Cummins et al. (1994, 1996) estimate capital's responsiveness by focusing on those periods with major tax reforms in order to reduce measurement error in the user cost. They report large σ's (but see note 2 to the table for an alternative interpretation that lowers these estimates). Also focusing on a major tax reform but using Mexican establishment panel data, Ramírez-Verdugo (2006) reports an elasticity of 1.10. Clark (1993) uses a smaller panel for 15 classes of equipment assets and reports σ's ranging from 0.25 to his preferred estimate of 0.40. These estimates translate into σ's for total capital of roughly 0.18–0.28. With a large panel of firm specific data, Chirinko et al. (1999) use a specification very similar to Eq. (7) modified to account for firm fixed effects, and they obtain a precisely estimated but small value for σ of 0.25.

The major concern with the implicit investment equation presented in Eq. (7) is similar to the one surrounding explicit models. Since both frameworks are based on short-run data that represents adjustments toward new steady-states, both are affected by frictions that affect the dynamic accumulation path. As listed at the end of Section 3.1 many frictions potentially affect the firm, and there is little assurance that they are fully captured by the distributed lags in Eq. (7a). The effect of misspecified dynamics on estimates of σ remains unclear.

4. σ: The long of It

Faced with difficulties modeling frictions and the associated dynamics, some researchers have turned to an alternative estimation strategy focusing on the long-run relations that appear in the static model. This strategy has the decided advantage of avoiding the numerous specification issues that confront models using short-run investment data. The challenge becomes how to deal with the unobservability of the long-run variables (denoted with a "*"). Three specific approaches are discussed, and each focuses on the first-order condition for capital (Eq. (4a)).

The Unadjusted First-Order Condition model estimates Eq. (4a) by assuming that the observed values of the capital/output ratio and user cost are reasonably close to their long-run values and hence distinctions between ky_t^* & ky_t and ucc_t^* & ucc_t can be relegated to the error term. Studies listed in panel C.1 have followed this approach, and the results have varied widely, though the estimated σ's tend to be above 0.50. The conference paper by Klump, McAdam, and Willman also adopts this approach. They expand on the Unadjusted First-Order Condition model by carefully modeling technical change and by estimating a system of equations – the first-order conditions for capital and labor and the production function. The authors' preferred estimate is $\sigma = 0.70$.

A second way to model unobservable long-run variables is to exploit Cointegration properties. In an innovative paper, Caballero (1994) measures long-run values by exploiting the Cointegration relations between the capital/output ratio and the user cost of capital. He assumes that the ky_t and ucc_t series are $I(1)$ and cointegrated; thus long-run movements in ky_t and uc_t dominate. (If the two series are $I(0)$, then the Cointegration model no longer measures long-run values.) The cointegrating vector is $(1, \sigma)$.[7] The Cointegration model generates consistent estimates in a large sample but, in finite samples, the estimates are biased by correlation between the user cost and the error term. This bias is accounted for with the Stock and Watson (1993) correction, which involves adding leads and lags of the difference of the right-side variable. This correction has a substantial influence on the estimated σ's in some applications and can be interpreted as controlling for deviations between long-run and observed values distorting parameter estimates. The cointergration model can be written as follows:

$$ky_{i,t} = \gamma - \sigma \text{ucc}_{i,t} + \sum_{j=1}^{J} \omega_j \Delta\text{ucc}_{i,t-j} + \sum_{j=1}^{J} \xi_j \Delta\text{ucc}_{i,t+j} + v_{i,t}^{ky}, \tag{9}$$

where the "i" subscripts on the variables represents the panel nature of the data, γ, σ, the ω's, and the ξ's are parameters to be estimated, and $v_{i,t}^{ky}$ is an error term that contains both a stochastic component and fixed effects.

Studies listed in panel C.2 have followed this approach, and the estimated σ's fall into two categories. Estimates based on aggregate data are 1.20 for Canada and a substantial 3.40 for Mexico. The latter result may be due to some characteristic of a developing economy or the focus on major tax reforms. Other studies based on US and UK data report sharply lower estimates. The original study by Caballero with aggregate US data yields an estimate of 0.65, which is similar to the estimate of 0.70 reported by Caballero et al. (1995) based on US panel data. Lower estimates ranging between 0.32 and 0.42 are obtained by Barnes et al. (2006) using a very similar approach on UK panel data. The conference paper by Smith also estimates a Cointegration model with UK panel data, and his preferred estimate is $\sigma = 0.40$.

The third approach deals with the long-run variables directly. As developed in Chirinko et al. (2007), this Interval-Difference model divides the sample period in half, averages the data in each interval to form long-run variables, and then differences the interval-averaged data to eliminate individual effects. The effect of technical change is captured in the resulting constant term. The Interval-Difference model is written as follows:

$$\Delta ky_i^{\#} = \chi - \sigma \Delta\text{ucc}_i^{\#} + \Delta v_i, \tag{10}$$

[7] As argued in Chirinko and Mallick (2007b), this estimation strategy faces some econometric difficulties in recovering production function parameters.

where the # superscript represents interval averages. In Eq. (10), the temporal dimension of the data has been eliminated by averaging within each of the two intervals and then first-differencing. The parameters are estimated in a cross-section regression.

This procedure may seem both appealing but ad hoc. However, the averaging procedure has a well-defined interpretation as a low-pass filter in the frequency domain. Filters are transformations – usually linear, two-sided (i.e., leads and lags), and symmetric – of time series that allow researchers to emphasize certain aspects of the data. Filters interpreted in the frequency domain allow researchers to emphasize certain frequencies (low, medium, or high) believed to be relevant for the problem at hand. Chirinko et al. (2007) allow long-run (low) frequencies to pass through the filter but exclude the medium and high frequencies that are inappropriate for production function estimation. They define the long-run as movements only at the 0th frequency (though the results are robust to including additional low frequencies in defining the long-run). Baxter and King (1999, Section II.B), present the formulas that allow frequency domain interpretations to be used with data defined in the time domain.[8] Frequency domain filters are best used on data that are stationary, though they can be appropriate with nonstationary data if the nonstationarity is induced by a trending mean or a factor that is trending slowly. The Interval-Difference model has been applied to US and UK panel data. As shown in panel C.3, the estimates are very close, ranging between 0.32 and 0.40.[9]

The results presented in the table range widely. Even when we focus on the models in Panels C.2 and C.3 that make explicit adjustments in order to emphasize long-run variation, a consensus value remains elusive. In particular, elasticities from the Cointegration model tend to be larger than those obtained from the Interval-Difference model. This result is particularly puzzling because the Cointegration and Interval-Difference models both emphasize long-run variation at the 0th frequency.

Apart from any inherent tendency toward generating high or low estimates of σ, the results reported above for the Cointegration and Interval-Difference models might differ, *inter alia*, because of the specification of the user cost of capital variable, the use of plant-level vs. firm-level data, or the sample period. To abstract from these sources of variation, we report the results from Chirinko et al. (2007, Table 5), who use the same panel dataset for US firms to estimate each of the three models considered in this section – the Unadjusted First-Order Condition (Eq. (4a) without the *'s), Cointegration (Eq. (9)), and Interval-Difference (Eq. (10)) models. In order to assure that the basis of comparison is valid, we begin by reporting a Cointegration model that is close to the specification used by Caballero et al. (1995) with three lags of Δucc_t, without time fixed effects, and with firm-level rather than plant-level data. The estimated $\sigma = 0.72$, which is very close to the estimate of 0.70 reported in the table.[10] The estimate rises to 0.75 when time fixed

[8] The Hodrick–Prescott filter is another candidate filter. Relative to the Baxter–King filter, it does not have a natural interpretation but, in principle, has the benefit of retaining data at the beginning and end of the samples. However, as noted by Baxter and King, it is advisable to delete the extreme endpoints from the Hodrick–Prescott filter; thus the gain with additional data is not realized in practice.

[9] The results reported by Chirinko and Mallick (2007a) are based on a variant of the Interval-Difference model in which data are averaged for each year in the sample (subject to data availability at the beginning and end of the sample). The averaged data are first-differenced but, since there are more than two intervals, the model is estimated as a panel (i.e., time subscripts are added to the variables in Eq. (10)).

[10] All estimates of σ reported in the remaining part of this section are statistically significant at conventional levels; the lowest t-statistic exceeds 6.0.

effects are included. Thus, the choice of firm-level or plant-level data does not affect the estimate of σ.

With this benchmark established, we are now in a position to compare the three models. The estimated σ is 0.52 for the Unadjusted First-Order Condition model, a decline of 30% from the benchmark value of $\sigma = 0.75$. A similar decline appears when the full Stock–Watson correction with three leads and lags (and time fixed effects) is applied to the Cointegration model.[11] In this case, σ equals 0.54. Thus, the Stock–Watson correction that matters in other applications of the Cointegration model is not important in the current case.

While there is no noticeable difference between estimates from the Unadjusted First-Order Condition and Cointegration models, there remains a statistically significant gap of 0.12–0.14 between these estimates and $\sigma = 0.40$ from the Interval-Difference model. We do not have a full resolution of this puzzling gap but offer two provisional explanations. The time-series properties of the data may be a contributing factor. One of the key assumptions underlying the consistency of the estimation strategy associated with the Unadjusted First-Order Condition and Cointegration models is that the model variables are non-stationary. Chirinko et al. (2007) test this assumption and reject the null hypothesis of non-stationarity for ucc_t. This result seems reasonable since equilibrating forces will prevent the user cost of capital from wandering-off too far. For example, if ucc_t rose dramatically, economic activity would slow and the price of investment goods and the rate of interest, both of which enter the user cost, would decline. A second explanation of the gap is that the source of identification is fundamentally different. The Unadjusted First-Order Condition and Cointegration models achieve identification by the time-series properties of the variables at the 0th frequency. By contrast, The Interval-Difference model relies on the cross-section variation in long-run growth rates of the capital/output ratio and the user cost to identify σ. The direction of bias associated with these two explanations is unclear. Nonetheless, the gap is small compared to the range of estimates in the table.

5. Conclusions and lingering issues

This review of models and results leads to several conclusions and some lingering issues worthy of further consideration:

A. The "long and short of it" is that long-run models would seem to be preferred for estimating production function parameters such as σ. A fundamental tension exists between the short-run data that are available and the long-run parameter that is required. Models focusing on long-run behavior have the decided advantage of avoiding the difficult problems associated with modeling frictions and the associated short-run dynamics that may blur estimates of σ.

B. While some estimates of σ are above one, the weight of the evidence suggests that σ lies in the range between 0.40 and 0.60.

[11] The full Stock–Watson correction was not implemented in the original Caballero et al. (1995) paper because of a concern for lost degrees of freedom and the results from Caballero (1994) that estimates of σ are robust to using only lagged correction terms.

Table 1
Estimates of the elasticity of substitution[a]

	σ	Characteristics of the study
A. Investment data – aggregate		
Jorgenson (1963)	1.00	Cobb–Douglas production function
Hall and Jorgenson (1967, 1971)	1.00	Cobb–Douglas production function
Eisner and Nadiri (1968)	0.16 to 0.33	CES production function
Chirinko (1993a,b)	0.00 to 0.30	Survey of econometric estimates
Tevlin and Whelan (2003)	0.18	US aggregate investment
	1.59	US computer investment
Bakhshi et al. (2003)	0.32	UK aggregate investment
	1.33	UK computer investment
Ellis and Price (2004)	0.44	UK aggregate investment
B. Investment data – panel		
Cummins and Hassett (1992)		Years of major tax reforms
Equipment	0.93 [0.23[b]]	
Structures	0.28 [0.07[b]]	
Clark (1993)	0.18 to 0.28[c]	Fifteen asset classes
Cummins et al. (1994, 1996)	0.67 [0.17[b]]	Years of major tax reforms
Chirinko et al. (1999)	0.25	Variety of estimators
Ramírez-Verdugo (2006)	1.10	Major Mexican tax reform
C. Capital stock data		
1. *Unadjusted first-order condition*		
Lucas (1969)	0.30 to 0.60	Variety of specifications
Berndt (1976)	0.00 to 1.24	Variety of specifications
Berndt (1991)	0.97	Translog system
Jorgenson and Yun (2001)		Translog system
Corporate	0.50	
Noncorporate	0.70	
Chirinko et al. (2007)	0.52	US panel data
Klump et al. (2007)	0.60	Three equation system, US data
Klump et al. (2008)	0.70	Three equation system, Euro Area data
2. *Cointegration model*		
Caballero (1994)	0.65[c]	US aggregate data
Caballero et al. (1995)	0.70[c]	US (plant) panel data
Schaller (2006)	1.20	Canadian aggregate data
Ramírez-Verdugo (2006)	3.40	Mexican panel data and major tax reforms
Barnes et al. (2006)	0.32 to 0.42	UK panel data
Chirinko et al. (2007)	0.54	US panel data
Smith (2008)	0.40	UK panel data
3. *Interval-difference model*		
Barnes et al. (2006)	0.32	UK panel data
Chirinko et al. (2007)	0.40	US panel data
Chirinko and Mallick (2007a)	0.33	US (industry) panel data

[a] This table is taken from Chirinko (2002, Table 1) and updated with several recent results. The results are presented in chronological order within a category.

[b] Chirinko et al. (1999, Section 5) offer a different interpretation that lowers the originally reported estimates to the figures in brackets because Cummins et al.'s econometric equation contains the level (rather than the percentage change) in user cost.

[c] The cited study presents elasticity estimates for equipment capital that are translated into a total capital elasticity by multiplying the reported elasticity by 0.70. This adjustment factor is obtained from the following calculation. Cummins and Hassett (1992) estimate separate elasticities for equipment and structures capital, and find that the equipment elasticity is larger by a factor of three. (An identical factor is obtained by Pindyck and Rotemberg (1983b).) Assuming that equipment and structures have (stock) weights of 0.55 and 0.45, respectively, the elasticity for equipment and structures is computed from the equipment estimate (0.40 in this example) as follows, $0.55 * 0.40 + 0.45 * 0.40/3 = 0.40 * 0.70 = 0.28$.

C. The evidence rather strongly rejects the Cobb–Douglas assumption of σ equal to one. While convenient analytically, the Cobb–Douglas assumption is inappropriate empirically. This assumption continues to be used in policy analysis (e.g., US Department of the Treasury, 2006), and hence the simulated impacts of policy changes may well be overstated.
D. As emphasized in several of the theoretical papers and documented in the Klump, McAdam, and Willman paper, biased technical progress is an important element that needs to be given careful consideration in production function estimation.
E. The substantial variation in panel data has several benefits to recommend it. While panel data may be preferred on econometric grounds, estimates based on disaggregate data lead to a further problem. What is the "mapping" between the disaggregate parameters that have been estimated and the aggregate parameter that is of interest? The results with panel data reported in Table 1 (with the exception of Chirinko and Mallick, 2007a; Smith, 2008) constrain σ to be the same across sectors. However, it would not seem unreasonable that σ differs among firms or across industries. In this case, how do we infer an aggregate σ from disaggregate estimates?[12] Such an issue has been on the research agenda for some time. In the "Speculation" sub-section concluding their paper, Arrow et al. (1961, p. 247) suggest that

> Given systematic intersectoral differences in the elasticity of substitution and in income elasticities of demand, the possibility arises that the process of economic development itself might shift the over-all elasticity of substitution.

Further work on such a "mapping" from disaggregate to aggregate parameters that accounts for the effects of intersectoral development would seem of much importance in empirical work on production functions and for the value of σ relevant for aggregate growth models.

Acknowledgements

The author gratefully acknowledges financial support from the Bank of England under a Houblon-Norman/George Senior Fellowship and comments by Zheng Liu, Peter McAdam Simon Price, and participants at Goethe University Frankfurt/CES Conference, *A Bright Future at the Age of 50 – The CES Production Function In the Theory and Empirics of Economic Growth*, especially Uwe Hassler, Rainer Klump, and James Smith. Special thanks is owed to Steven Fazzari, Debdulal Mallick, and Andrew Meyer; several of the perspectives and results presented in this paper have been developed over the years in our joint work. All errors and omissions remain the sole responsibility of the author, and the conclusions are not necessarily shared by the organizations with which he is associated.

References

Abel, A.B., Blanchard, O.J., 1986. The present value of profits and cyclical movements in investment. Econometrica 54 (March), 249–274.

[12] This aggregation issue has been addressed in the recent work of Miyagiwa and Papageorgiou (2007) and Palivos (2008).

Akerlof, G., Nordhaus, W.D., 1967. Balanced growth – A razor's edge? International Economic Review 8 (October), 343–348.

Arrow, K.J., Chenery, H.B., Minhas, B.S., Solow, R.M., 1961. Capital-labor substitution and economic efficiency. The Review of Economics and Statistics 43, 225–250 (reprinted in production and capital: Collected papers of Arrow, K.J., vol. 5. Cambridge: Harvard University Press, 1985, pp. 50–103).

Bakhshi, Hasan, Oulton, Nicholas, Thompson, Jamie, 2003. Modelling investment when relative prices are trending: Theory and evidence for the United Kingdom. Bank of England Working Paper No. 189.

Barnes, Sebastian, Sebastia-Barriel, Maria, Price, Simon, 2006. The elasticity of substitution: Evidence from a UK Firm-Level Data Set. Bank Of England, September.

Barro, R.J., Sala-i-Martin, X., 1995. Economic Growth. McGraw-Hill, New York.

Baxter, M., King, R.G., 1999. Measuring business cycles: Approximate band-pass filters for economic time series. The Review of Economics and Statistics 81 (November), 575–593.

Berndt, E.R., 1976. Reconciling alternative estimates of the elasticity of substitution. The Review of Economics and Statistics 63 (February), 59–68.

Berndt, E.R., 1991. The Practice of Econometrics: Classic and Contemporary. Addison-Wesley, Reading, MA.

Blackorby, C., Russell, R.R., 1989. Will the elasticity of substitution please stand up? (A comparison of the Allen/Uzawa and Morishima elasticities). American Economic Review 79 (September), 882–888.

Caballero, R.J., 1994. Small sample bias and adjustment costs. The Review of Economics and Statistics 76 (February), 52–58.

Caballero, R.J., 1999. Aggregate investment. In: Taylor, J.B., Woodford, M. (Eds.), . In: Handbook of Macroeconomics, vol. 1B. Elsevier North-Holland, Amsterdam, pp. 13–862.

Caballero, R.J., Engel, E.M.R.A., Haltiwanger, J.C., 1995. Plant-level adjustment and aggregate investment dynamics. Brookings Papers on Economic Activity (2), 1–54.

Caselli, F., 2005. Accounting for cross-country income differences. In: Aghion, P., Durlauf, S. (Eds.), Handbook of Economic Growth, vol. 1A. Elsevier North-Holland, Amsterdam, pp. 679–741.

Chirinko, R.S., 1993a. Business fixed investment spending: Modeling strategies, empirical results, and policy implications. Journal of Economic Literature 31 (December), 1875–1911.

Chirinko, R.S., 1993b. Econometric Models and Empirical Findings for Business Investment. Basil Blackwell, New York.

Chirinko, R.S., 2002. Corporate taxation, capital formation, and the substitution elasticity between labor and capital. National Tax Journal 55 (June), 339–355.

Chirinko, R.S., Mallick, D., 2007a. The substitution elasticity, growth theory, and the low-pass filter panel model. Emory University, January.

Chirinko, R.S., Mallick, D., 2007b. The Fisher/Cobb–Douglas paradox, factor shares, and cointegration. Emory University, June.

Chirinko, R.S., Fazzari, S.M., Meyer, A.P., 1999. How responsive is business capital formation to its user cost? An exploration with micro data. Journal of Public Economics 74 (October), 53–80.

Chirinko, R.S., Fazzari, S.M., Meyer, A.P., 2007. That elusive elasticity: A long-panel approach to estimating the capital-labor substitution elasticity. Emory University, June.

Clark, P.K., 1993. Tax incentives and equipment investment. Brookings Papers on Economic Activity (1), 317–347.

Cummins, J.G., Hassett, K.A., 1992. The effects of taxation on investment: New evidence from firm level panel data. National Tax Journal 45 (September), 243–252.

Cummins, J.G., Hassett, K.A., Hubbard, R.G., 1994. A reconsideration of investment behavior using tax reforms as natural experiments. Brookings Papers on Economic Activity (2), 1–60.

Cummins, J.G., Hassett, K.A., Hubbard, R.G., 1996. Tax reforms and investment: A cross-country comparison. Journal of Public Economics 62 (October), 237–273.

de La Grandville, O., 1989. In quest of the slutsky diamond. American Economic Review 79 (June), 468–481.

de La Grandville, O., 1997. Curvature and the elasticity of substitution: Straightening it out. Journal of Economics 66 (February), 23–34.

de La Grandville, O., Solow, R.M., 2004. On the Determinants of Economic Growth: Is Something Missing? University of Geneva and MIT.

Diamond, P., McFadden, D., Rodriguez, M., 1978. Measurement of the elasticity of factor substitution and bias of technical change. In: Fuss, M., McFadden, D. (Eds.), Production Economics: A Dual Approach to Theory and Applications, vol. 2. North-Holland, Amsterdam, pp. 125–147.

Duffy, J., Papageorgiou, C., 2000. A cross-country empirical investigation of the aggregate production function specification. Journal of Economic Growth, 87–120.
Eisner, R., Nadiri, M.I., 1968. Investment behavior and neo-classical theory. The Review of Economics and Statistics 50 (August).
Eisner, R., Strotz, R.H., 1963. Determinants of business investment. In: Commission on Money and Credit, Impacts of Monetary Policy. Prentice-Hall, Englewood Cliffs, NJ, pp. 60–337.
Ellis, C., Price, S., 2004. UK business investment and the user cost of capital. Manchester School 72 (September), 72–93.
Engen, E., Gravelle, J., Smetters, K., 1997. Dynamic tax models: Why they do the things they do. National Tax Journal 50 (September), 657–682.
Farmer, R.E.A., 1997. Money in a real business cycle model. Journal of Money, Credit, and Banking 29 (Part 2), 568–611.
Fazzari, S.M., Hubbard, R.G., Petersen, B.C., 1988. Financing constraints and corporate investment. Brookings Papers on Economic Activity (1), 141–195.
Felipe, J., Holz, C.A., 2001. Why do aggregate production functions work? Fisher's simulations, Shaikh's identity, and some new results. International Review of Applied Economics 15, 261–285.
Hall, R.E., Jorgenson, D.W., 1967. Tax policy and investment behavior. American Economic Review 57 (June), 391–414.
Hall, R.E., Jorgenson, D.W., 1971. Application of the theory of optimum capital accumulation. In: Fromm, G. (Ed.), Tax Incentives and Capital Spending. Brookings Institution, Washington, pp. 9–60.
Hamermesh, D., 1993. Labor Demand. Princeton University Press, Princeton.
Hicks, J.R., 1963. The Theory of Wages, second ed. MacMillan & Co., London (first edition published in 1932).
Hubbard, R.G., 1998. Capital-market imperfections and investment. Journal of Economic Literature 36 (March), 193–225.
Jorgenson, D.W., 1963. Capital theory and investment behavior. American Economic Review 53 (May), 247–259.
Jorgenson, D.W., Yun, K.-Y., 2001. Lifting The Burden: Tax Reform, The Cost of Capital, and US Economic Growth. MIT Press, Cambridge.
King, R.G., Rebelo, S.T., 1993. Transitional dynamics and economic growth in the neoclassical model. American Economic Review 83 (September), 908–932.
Klump, R., de La Grandville, O., 2000. Economic growth and the elasticity of substitution: two theorems and some suggestions. American Economic Review 90 (March), 282–291.
Klump, R., Preissler, H., 2000. CES production functions and economic growth. Scandinavian Journal of Economics 102 (March), 41–56.
Klump, R., Saam, M., 2007. Calibration of Normalized CES Production Functions in Dynamic Models. Goethe University Frankfurt.
Klump, R., McAdam, P., Willman, A., 2007. Factor substitution and factor augmenting technical progress in the US: A normalized supply-side system approach. The Review of Economics and Statistics 89 (February), 183–192 (a preliminary version is published as ECB Working Paper No. 367, June 2004).
Klump, R., McAdam, P., Willman, A., 2008. Unwrapping some Euro Area Growth Puzzles: Factor Substitution, Productivity and Unemployment. Journal of Macroeconomics 30 (2), 645–666.
Lau, L.J., 1980. On the uniqueness of the representation of commodity-augmenting technical change. In: Klein, L.R., Marc, N., Sho-Chieh, T. (Eds.), Quantitative Economics and Development: Essays in Memory of Ta-Chung Liu. Academic Press, New York, pp. 281–290.
Lucas, R.E., 1969. Labor-capital substitution in US manufacturing. In: Harberger, A.C. (Ed.), The Taxation of Income from Capital. Brookings Institution, Washington, pp. 223–274.
Mallick, D., 2006. Reconciling Different Estimates of the Elasticity of Substitution. Emory University.
Mankiw, N.G., 1995. The growth of nations. Brookings Papers on Economic Activity (1), 275–310.
Miyagiwa, K., Papageorgiou, C., 2007. Endogenous aggregate elasticity of substitution. Journal of Economic Dynamics and Control 31 (September), 2899–2919.
Palivos, T., 2008. Comment on σ: The long and short of it. Journal of Macroeconomics 30 (2), 687–690.
Pindyck, R.S., Rotemberg, J.J., 1983a. Dynamic factor demands and the effects of energy price shocks. American Economic Review 73 (December), 1066–1079.
Pindyck, R.S., Rotemberg, J.J., 1983b. Dynamic factor demands under rational expectations. Scandanvian Journal of Economics 85 (3), 223–238.
Pitchford, J.D., 1960. Growth and the elasticity of substitution. Economic Record 36 (December), 491–504.
Ramírez-Verdugo, 2006. Tax Incentives and Business Investment. MIT.

Robinson, J., 1933. The Economics of Imperfect Competition. MacMillan & Co., London (reprinted 1959).

Schaller, H., 2006. Estimating the long-run user cost elasticity. Journal of Monetary Economics 53 (May), 725–736.

Schefold, Bertram, 2008. C.E.S. production functions in the light of the Cambridge critique. Journal of Macroeconomics 30 (2), 783–797.

Shaikh, A., 1974. Laws of production and laws of algebra: The humbug production function. Review of Economics and Statistics 56 (February), 115–120.

Smith, J., 2008. That elusive elasticity and the ubiquitous bias: Is panel data a panacea? Journal of Macroeconomics 30 (2), 760–779.

Solow, R.M., 1956. A contribution to the theory of economic growth. Quarterly Journal of Economics 70, 65–94.

Solow, R.M., 1974. Laws of production and laws of algebra: The humbug production function: A comment. Review of Economics and Statistics 56 (February), 121.

Solow, R.M., 1987. Second thoughts on growth theory. In: Weiserbs, D., Steinherr, A. (Eds.), Employment and Growth: Issues for the 1980s. Kluwer Academic Publishers, Dordrecht and Boston, pp. 13–28.

Stock, J.H., Watson, M.W., 1993. A simple MLE of cointegrating vectors in higher order integrated systems. Econometrica 61 (July), 783–820.

Tevlin, S., Whelan, K., 2003. Explaining the investment boom of the 1990s. Journal of Money, Credit and Banking 35 (February), 1–22.

Turnovsky, S.J., 2008. The role of factor substitution in the theory of economic growth and income distribution: Two examples. Journal of Macroeconomics 30 (2), 604–629.

US Department of the Treasury., 2006. Office of Tax Analysis, A Dynamic Analysis of Permanent Extension of the President's Tax Relief, US Department of the Treasury, 25th July.