The Substitution Elasticity, Growth Theory, And The Low-Pass Filter Panel Model

Robert S. Chirinko
Emory University and CESifo

and

Debdulal Mallick*
Emory University

January 2007

Preliminary -- Please Do Not Quote Nor Cite

* We would like to acknowledge the helpful comments and assistance offered by Stephen Bond, Hashem Dezhbakhsh, John Fernald, Simon Gilchrist, Nazrul Islam, Christian Murray, Franz Palm, Kevin Stiroh, John van Reenen, Dan Wilson, Mahmut Yasar and by participants in Emory's Macro Lunch Group and at seminars at the University of Frankfurt, the 11th International Conference On Panel Data, the London School of Economics /Institute of Fiscal Studies/University College London (joint seminar), the 2006 European Economic Association Meetings, and the 2006 NBER Summer Institute (Productivity/Macroeconomics), especially our discussant, Ernst Berndt. We are particularly grateful to Elena Pesavento for her thoughtful comments on a prior draft. Dale Jorgenson has generously made his dataset available to us. Financial assistance from Emory University's ICIS Research and Program Fund and the Graduate School is gratefully acknowledged. All errors, omissions, and conclusions remain the sole responsibility of the authors.

The Substitution Elasticity, Growth Theory, And The Low-Pass Filter Panel Model

Abstract

The elasticity of substitution between labor and capital (σ) is a crucial parameter in growth theory. A host of important issues, including the possibility of perpetual growth or decline, depend on the precise value of σ. This paper examines the role of σ in the neoclassical growth model and estimates σ by combining a low-pass filter with standard panel data techniques to identify the long-run relations appropriate to production function estimation. Our approach is in the spirit of Friedman's permanent income theory of consumption and Eisner's related permanent income theory of investment. While their approaches and ours are similar in relying on permanent components, we extract these components with spectral methods that are more powerful and general for identifying these unobservables. We transform the data with the Baxter-King low-pass filter that depends on two parameters, the critical periodicity defining the long-run frequencies and a window for the number of lags approximating the ideal low-pass filter. Based on an analysis of the spectrum of the transformed series, we confirm that our choices of the critical periodicity and window emphasize long-run variation.

The empirical results are based on the comprehensive panel industry data constructed by Dale Jorgenson and his research associates. Our preferred estimate of σ is 0.30 for the benchmark values of the critical periodicity of eight years and window of three years. This result is robust to variations in the window. As the periodicity declines from eight to the minimum value of two, the elasticity declines owing to the distorting effects of transitory variation.

JEL Nos.: O4, C23

Keywords: Substitution elasticity, growth theory, panel data, low-pass filter

Corresponding Author:
Robert S. Chirinko
Department of Economics
Emory University
Atlanta, Georgia 30322-2240
U.S.A.

Debdulal Mallick
Department of Economics
Emory University
Atlanta, Georgia 30322-2240
U.S.A.

EM: robert.chirinko@emory.edu, PH: 1-404-727-6645, FX: 1-404-727-4639

Table Of Contents

The Substitution Elasticity, Growth Theory, And The Low-Pass Filter Panel Model

> A "crucial" assumption is one on which the conclusions do depend sensitively, and it is important that crucial assumptions be reasonably realistic.
>
> Solow (1956, p. 65)

> The relation between the theoretical constructs used in consumption research and the observable magnitudes regarded as approximating them has, I believe, received inadequate attention.
>
> Friedman (1957, p. 7)

I. Introduction

The essential innovation contained in the neoclassical growth model was the modification of the steady-state equilibrium condition. Prior to Solow (1956), growth models determined the steady-state with a set of independent parameters, and equilibrium was achieved in only the most unlikely of circumstances. Solow's innovation was to introduce a variable capital/output ratio in place of the Harrod-Domar fixed parameter. This tour de force resolved the pressing analytic problem of the knife-edge solution inherent with independent parameters by relaxing the "crucial" assumption that production occurs with fixed factor proportions. However, this innovation focused attention onto other "crucial" assumptions embedded in the neoclassical production function. Key among these is the elasticity of substitution between labor and capital, σ, a parameter that has received too little notice in the growth literature.

This paper examines the role of σ in the neoclassical growth model and estimates this parameter by combining a low-pass filter with standard panel data techniques to identify the long-run relations appropriate to production function estimation. We begin in Section II with a discussion of the pivotal role played by σ in a variety of issues in growth theory -- the possibility of perpetual growth or decline, the level of steady-state income per capital, the speed of convergence, the rate of return on capital, the role of biased technical change, and the allocation of per capita income between factors of production and the efficiency with which they are utilized. In one sense, Solow's fundamental innovation replaces the Harrod-Domar assumption that $\sigma = 0$ with the more general assumption that $\sigma \geq 0$. Our discussion highlights that a host of important growth issues depend on the precise positive value of σ.

Section III develops a strategy for estimating σ's that apply a low-pass filter based on spectral methods to panel data.[1] Production function parameters are recovered by focusing on the long-run relations between arguments appearing in the first-order condition for capital. Our approach is in the spirit of Friedman's (1957) permanent income theory of consumption and Eisner's (1967) related permanent income theory of investment. Friedman observed that the fundamental relation between consumption and income obtained between their permanent components and then identified the permanent components in terms of geometric distributed lags of past values. Eisner also emphasized the distinction between the transitory and permanent components of variables affecting investment demand. He isolated the permanent component by grouping firms by industry and then estimating with the group means. While our approach also relies on permanent components, they are extracted with spectral methods that are more powerful and

[1] To the best of our knowledge, the only other study using spectral methods to study capital formation is Engle and Foley (1975), who estimate a model relating investment spending to an equity price series (approximately a Brainard-Tobin's Q variable) and use a band-pass filter to emphasize middle frequencies centered at two years.

general for identifying these unobservable variables. We transform the data with a low-pass filter developed by Baxter and King (1999) that depends on two parameters, the critical periodicity ($p^{\#}$) defining long-run frequencies and a window (q) for the number of lags and leads used to approximate the ideal low-pass filter.

Section IV examines the theoretical properties of the spectral representation of the low-pass filter. The spectra associated with our estimator are derived and used to assess the extent to which our choices of the critical periodicity and window are successful in emphasizing long-run variation. We vary the key periodicity ($p^{\#}$) and window (q) parameters to assess the sensitivity of the allocation of variance across low, middle, and high frequencies and verify that the transformed data reflect long-run variation that closely matches production function concepts.

Section V to VII contain empirical results based on the comprehensive panel of U.S. industry data constructed by Dale Jorgenson and his research associates. Our econometric model relates the long-run capital/output ratio to the long-run relative price of capital. The benchmark estimate of σ is 0.30 for the critical periodicity of eight years ($p^{\#} = 8$) and a window of three years ($q = 3$). This result is robust to variations in the window. Moreover, there is great value in using the spectral methods to extract permanent components. As the periodicity declines from eight to the minimum value of two, the estimated σ declines by one-third owing to the distorting effects of transitory variation. At $p^{\#} = 2$, the low-pass filter is neutral, the raw data are not transformed, low frequency variation is not emphasized, and σ reaches a lower bound (relative to other values of $p^{\#}$) of 0.21.

Section VI contains three robustness checks. First, the benchmark estimates are based on the assumptions that measurement error or simultaneity is absent, and hence OLS is the appropriate estimation technique. Instrumental variables estimates confirm the prior OLS results. We also verify that the instruments are relevant. Second, the benchmark estimates are based on the assumption of a constant σ across time; we document that our estimate of σ is temporally stable. Third, we examine alternative specifications for estimating σ. The first-order conditions yield two additional estimating equations that contain the labor/output ratio or the labor/capital ratio as the dependent variable and a relative price multiplying σ. However, the neoclassical growth model implies that neither series is stationary, a prediction consistent with the stylized facts of growth. Hence, the low-pass filter used in this study is not applicable because spectral methods require stationary data. We nonetheless examine the estimates of σ derived from these models given the important study of Berndt (1976) that reported a disturbing wide range of estimates. Berndt's findings about the relation among σ's estimated from the three different models are confirmed.

Section VII contains results when the σ's are allowed to vary by industry. Given our interest in the impact of aggregate σ on growth theory, the homogeneity assumption imposed across industries in the prior two sections is a natural way of obtaining the substitution elasticity. From an estimation perspective, however, it might be desirable to exploit the panel feature of our dataset and to allow the σ's to differ across industries. Since ultimate interest resides with an aggregate substitution elasticity, we develop a mapping from the estimated industry parameters to the aggregate parameter of interest for growth theory. This mapping recognizes substitution within an industry and reallocations across industries. Additionally, we show that, if factor shares are independent of demand elasticities at the industry level, then reallocations vanish and the aggregate σ is a simple weighted-average of the industry σ's (with an additional minor adjustment). We find that the aggregate σ from the heterogeneous industry σ model is economically similar to that from the homogeneous σ model. The results for both homogeneous and heterogeneous models confirm that σ is far below the Cobb-Douglas value of unity.

Section VIII summarizes and concludes.

II. The Implications of σ for Growth Theory

The elasticity of substitution (σ) was introduced by Hicks (1932) to analyze changes in the shares of capital and labor. 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 (Blackorby and Russell, 1989, p. 882). It is well known that σ is important in, among other areas, the analysis of trade and factor returns (Jones and Ruffin, 2003) and tax policies (Chirinko, 2002). Less well known is the pivotal role played by σ in models of economic growth. Several prominent issues and their dependence on the value of σ are discussed in this section.

II.A. From The Harrod-Domar Knife-Edge to the Solow Interval

The neoclassical revolution in growth theory places the burden of equilibrium on the properties of the production function. When σ equals unity, the capital/labor ratio (K/L) converges to a positive, finite value because, as K/L moves towards its limiting values of 0 or ∞, the marginal product of capital (MPK[K/L : σ]) and average product of capital (APK[K/L : σ]) both tend to ∞ or 0, respectively. Thus the Inada conditions are satisfied and, as determined by Solow's fundamental equation of motion for K/L, capital accumulation converges to zero. However, when σ departs from unity, some interesting possibilities arise, and the capital stock and per capita income can exhibit perpetual decline or perpetual growth. Whether these outcomes obtain depends on the relation of σ to two critical values that depend on other parameters of the neoclassical growth model. These relations are portrayed in Figure 1. Values of σ greater than (less than) a critical value, $\sigma_H^{\#}$ ($\sigma_L^{\#}$), lead to perpetual growth (perpetual decline) in the capital stock. More standard behavior occurs for values of σ between $\sigma_H^{\#}$ and $\sigma_L^{\#}$; in this no growth case, capital accumulation converges to zero and K/L to a positive, finite value. The neoclassical growth model replaces the Harrod-Domar knife-edge with the Solow interval defined by the two critical σ's.

To examine the role of σ in generating non-standard equilbria, we need to consider the Solow's equation of motion for K/L and the limiting behavior of the average product of capital.[2] The well-known equation of motion in the neoclassical growth model is as follows,

$$\frac{\dot{\overline{K/L}}}{K/L} = s * APK[K/L;\sigma] - (n+\delta), \tag{1}$$

where s, n, and δ are the rates of saving, population growth, and depreciation, respectively. The APK[K/L : σ] is derived from the following intensive form of the CES production function (ignoring in this section the role of technical change),

$$f[K/L:\sigma] = \{\phi(K/L)^{((\sigma-1)/\sigma)} + (1-\phi)\}^{(\sigma/(\sigma-1))} \qquad 0 \le \phi \le 1, \tag{2}$$

where f[K/L:σ] is a per capita neoclassical production function depending on σ and ϕ, the capital distribution parameter. The MPK[K/L: σ] and APK[K/L: σ] are as follows,

[2] The analysis is this sub-section draws on the presentations of the neoclassical growth model in Barro and Sala-i-Martin (1995, Section 1.3.3), de La Grandville (1989), Klump and Preissler (2000), Klump and de La Grandville (2000), and de La Grandville and Solow (2004). The latter paper also discusses how increases in σ expand production possibilities in a manner similar to exogenous technical progress.

$$\text{MPK}[K/L:\sigma] \equiv f_k[K/L:\sigma] = \phi\{\phi + (1-\phi)(K/L)^{((1-\sigma)/\sigma)}\}^{(1/(\sigma-1))}, \quad (3)$$

$$\text{APK}[K/L:\sigma] \equiv \frac{f[K/L:\sigma]}{K/L} = \{\phi + (1-\phi)(K/L)^{((1-\sigma)/\sigma)}\}^{(\sigma/(\sigma-1))}. \quad (4)$$

The two non-standard cases arise because 1) MPK[K/L: σ] fails to satisfy one of the Inada conditions and 2) this positive, finite limit affects the APK[K/L: σ] so that no root exists for equation (1). In the first case where $\sigma > 1$, the limits for MPK and APK as K/L → ∞ are as follows,

$$\lim_{K/L\to\infty} \text{MPK}[K/L:\sigma] = \lim_{K/L\to\infty} \text{APK}[K/L:\sigma] = \phi^{(\sigma/(\sigma-1))}. \qquad \sigma > 1 \quad (5)$$

Perpetual growth arises when the limit in equation (5) is above the value of the APK[K/L: σ] that sets $((\dot{K/L})/(K/L) = 0)$ in equation (1). This critical value of the APK[K/L: σ] can be stated in terms of a critical high value of σ, $\sigma_H^{\#}$, determined by setting equation (1) to zero and solving for $\sigma_H^{\#}$ in terms of four other model parameters collected in $\Gamma_H = \{\delta, \phi, n, s_H : s_H > n + \delta\}$,

$$\begin{aligned} \sigma > \sigma_H^{\#} &\equiv g[\Gamma_H] \equiv g[\delta, \phi, n, s_H : s_H\phi > n + \delta] \\ &= \log[s_H/(n+\delta)]/\log[(\phi s_H)/(n+\delta)] > 1, \end{aligned} \quad (6)$$

where Γ_H represents the collection of four parameters such that $g[\Gamma_H] > 1$. When σ is high and substitution is relatively easy, the decrement to the marginal and average products of capital is moderate as K/L goes to infinity. If σ exceeds the critical value defined in equation (6), perpetual accumulation of capital and perpetual growth in per capita income are possible even in the absence of technical change. While receiving some sporadic attention over the past 50 years,[3] the distinct possibility of perpetual growth in the neoclassical model has received insufficient attention, being eclipsed by the popularity of endogenous growth models.

To gain some intuition for this result, note that the limits in equation (5) are increasing in any positive, finite value of σ. The higher is σ, the greater the "similarity" between capital and labor in the production function (Brown, 1968, p. 50). Assume that the increase in the capital/labor ratio represents an increment to capital with labor held fixed. When σ is high, the incremental capital is easily substituted for labor, resulting in a nearly equiproportionate increase in both factors. (The model thus begins to resemble an AK model popular in the endogenous growth literature.) Under constant returns to scale, diminishing returns sets-in very slowly, and the marginal and average products of capital can remain above the critical value so that capital accumulation is always positive.

In the second case when σ < 1, the other Inada condition fails,

$$\lim_{K/L \to 0} MPK[K/L : \sigma] = \lim_{K/L \to 0} APK[K/L : \sigma] = \phi^{(\sigma/(\sigma-1))} \qquad \sigma < 1, \qquad (7)$$

and this limit is below a low critical value,

$$\begin{aligned} \sigma < \sigma_L^{\#} &\equiv g[\Gamma_L] \equiv g[\delta, \varphi, n, s_L : s_L < n + \delta] \\ &= \log[s_L/(n+\delta)] / \log[(\phi s_L)/(n+\delta)] < 1. \end{aligned} \qquad (8)$$

[3] Solow (1956, pp. 77-78), Pitchford (1960), and Akerlof and Nordhaus (1967) were the first to note the possibility of perpetual growth. See the papers cited in fn. 2 for more recent statements.

When σ is low, capital and labor are “dissimilar” productive factors. With limited substitution possibilities, reductions in capital have little positive impact on marginal productivity. In an effort to raise the marginal product, capital accumulation remains negative and, for a value of σ below the $\sigma_L^{\#}$ defined in equation (8), K/L declines perpetually.

In their textbook on economic growth, Burmeister and Dobell (1970, p. 34) refer to situations where σ ≠ 1 as "troublesome cases" because they do not yield balanced growth paths. It is far from clear why the requirements for balanced growth paths in a particular theoretical model should dictate the shape of the production function, especially when σ = 1 is a sufficient but not necessary condition for a balanced growth path (cf. Acemoglu (2003) discussed below). To treat σ as a free parameter determined by the theory runs dangerously close to the fallacy of affirming the consequent. An alternative approach interprets cases where σ ≠ 1 as quite interesting, suggesting needed modifications to the standard neoclassical growth model and highlighting the key role played by σ.

II.B. Per Capita Income

The value of σ is linked to per capita income and growth. Klump and de La Grandville (2000) show that, for two countries with identical initial conditions (in terms of K/L, n, and s), the country with a higher value of σ experiences higher per capita income at any stage of development, including the steady state (if it exists).[4] De La Grandville (1989) argues theoretically that a relative price change (e.g., a decrease in the price of capital) leads to relatively more output the higher the value of σ. (He also notes a second channel depending on σ -- a higher substitution

[4] Miyagiwa and Papageorgiou (2003) demonstrate, however, that a monotonic relationship between σ and growth does not exist in the Diamond overlapping-generations model.

elasticity permits a greater flow of resources between sectors with different factor intensities.) Yuhn (1991) and Mallick (2006b) find empirical support of this hypothesis for South Korea and across 90 countries, respectively.

II.C. Speed of Convergence

The speed of convergence to the steady-state depends on σ through capital accumulation. Turnovsky's (2002, pp. 1776-1777) calibrated neoclassical growth model indicates that the rate of convergence is sensitive to and decreasing in σ. For a given productivity shock, the speed of convergence is 45.3% (per year) when σ equals 0.1, but drops markedly to 12.2% when σ equals 0.8. The speed of convergence falls further to 8.9%, 6.4%, and 3.5% as σ is increased to 1.0, 1.2, and 1.5, respectively.[5]

Several papers have shown that the influence of σ on the speed of convergence interacts with other parameters in the model. In Ramanathan (1975), the speed of convergence is negatively related to the share of capital. The larger the capital share, the less rapidly the APK declines and, since the APK is positively related to σ (cf. equation (4)), larger values of σ slow convergence. Mankiw (1995, p. 291) reports that an increase in the capital share from one-third to two-thirds reduces the speed of convergence by one-half. In the Klump and Preissler (2000, p. 50) model, the Ramanathan/Mankiw result holds, and the speed of convergence also depends on the relation between the initial and steady-state capital intensities.

[5] These figures are based on an intratemporal elasticity of substitution between consumption and leisure of 1.0 and an intertemporal elasticity of substitution for the composite consumption good of 0.4. The pattern of results is robust to variations in the latter parameter.

II.D. Other Issues in Growth Models

The value of σ can play an important role in assessing the plausibility of the neoclassical growth model. King and Rebelo (1993, Section IV) show that, in a Cass-Koopmans model with endogenous saving, the rate of return on capital (R) is sensitive to σ and is implausibly high when some part of growth is due to transitional dynamics. When transitional dynamics account for 25% of growth, R decreases moderately with σ. However, when transitional dynamics are more important, R increases dramatically with σ. Mankiw (1995, p. 287) also investigates the relation between σ and R but in terms of the following formula,

$$\frac{dR}{R} = \frac{-(1-\mu^K)}{\mu^K \sigma} * \frac{d(Y/L)}{Y/L}, \tag{9}$$

where μ^K is capital's factor share and (Y/L) is per capita income. Equation (9) approximates the difference in rates of return between poor and rich countries with the income differential represented by $\left(d(Y/L)/(Y/L)\right)$. For example, if σ = 4.0, the difference in the rate of return is only about 3 percent.[6] But if σ falls to 1.0 or 0.5, the above differential becomes implausibly large, increasing to 100 and 10,000 respectively. The relation between σ and R appears to be model dependent, but extant results suggest that the neoclassical model may not be correctly specified.

Purported differences in the rate of return to capital across countries has posed a persistent puzzle in the growth literature (Lucas, 1990). Recent work by Caselli and Feyrer (forthcoming) introduces a new method for calculating rates of return and shows that the cross-country dispersion in rates of return is substantially

---

[6] These computations are based on $\mu^K = 0.33$ and an income level in rich countries that is 10 times larger than in poor countries.

narrowed relative to prior estimates. While the estimates from this simple framework are independent of σ, the counterfactual exercises depend on production function characteristics. A value of $\sigma < 1$ is likely to reinforce the paper's general conclusions about the modest welfare benefits of reallocating the world's capital stock to enhance efficiency and the likely ineffectiveness of foreign aid to poorer countries.

The importance of technical change in growth models is sensitive to σ. Acemoglu (2003) examines the tension between fluctuations in income shares, the value of σ, and balanced growth. He develops a model in which technical change is both labor-augmenting and capital-augmenting and shows that, along the balanced growth path, all technical change will be labor-augmenting. If $\sigma < 1$, technical change stabilizes income shares, and the balanced growth path is stable and unique. In his review of developmental accounting (which assesses how much of cross-country differences in per capita income are attributable to productive factors and technical efficiency), Caselli (2005, Section 7) shows that the relative roles are very sensitive to σ. When σ is near 0.5, variation in productive factors accounts for almost 100% of the variation in per capita income across countries. The percentage is decreasing in σ and drops to 40% for $\sigma = 1.0$ (the Cobb-Douglas case) and 25% for $\sigma = 1.5$. Caselli (2005, end of Section 7) concludes that "gathering more information on this elasticity is a high priority for development accounting."

III. Estimation Strategy

III.A. The First-Order Condition

Our approach focuses on long-run production relations and low-frequency variation in the model variables. The long-run is defined by the vector of output and inputs consistent with profit maximization when all inputs can be adjusted without incurring costly frictions. This focus allows us to ignore short-run adjustment issues that are difficult to model and may bias estimates. Production for industry i at time t is characterized by the following Constant Elasticity of Substitution (CES) technology that depends on long-run values denoted by *,

$$Y^*_{i,t} = Y[K^*_{i,t}, L^*_{i,t}, A_{i,t}, B^K_t, B^L_t] \tag{10}$$
$$= A_{i,t}\left\{\phi(B^K_t K^*_{i,t})^{[(\sigma-1)/\sigma]} + (1-\phi)(B^L_t L^*_{i,t})^{[(\sigma-1)/\sigma]}\right\}^{[\sigma/(\sigma-1)]},$$

where $Y^*_{i,t}$ is long-run real output, $K^*_{i,t}$ is the long-run real capital stock, $L^*_{i,t}$ is the long-run level of labor input, ϕ is the capital distribution parameter, and σ is the elasticity of substitution between labor and capital. Technical progress is both neutral ($A_{i,t}$), and biased for capital and labor (B^K_t and B^L_t, respectively). Neutral technical change can have both industry and aggregate effects, and biased technical change, since it affects capital goods available to all industries, has an aggregate effect. Equation (10) is homogeneous of degree one in $K^*_{i,t}$ and $L^*_{i,t}$ and has three desirable features for the purposes of this study. First, this production function depends on only two parameters -- ϕ representing the distribution of factor returns and, most importantly, σ representing substitution possibilities between the factors of production. Second, the CES function is strongly separable and thus can be expanded to include many additional factors of production (e.g., intangible capital) without affecting the estimating equation derived below. This feature gives the CES specification an important advantage relative to other production functions that allow for a more general pattern of substitution possibilities (e.g., the translog, minflex-Laurent). Third, the Cobb-Douglas production function is a special case of the CES; as σ →1 and biased technical change disappears ($B_t^K = 1 = B_t^L$,), equation (10) becomes $Y_{i,t}^* = A_{i,t} \left\{ K_{i,t}^{*\,[\phi]} \, L_{i,t}^{*\,[1-\phi]} \right\}$.

Constrained by the CES production function (10), a profit-maximizing firm chooses capital so that its marginal product equals the Jorgensonian user cost of capital, defined as the price of capital, $P_{i,t}^{K*}$ (which combines interest, depreciation, and tax rates and the nominal price of capital goods), divided by the price of output, $P_{i,t}^{Y*}$. (The firm also sets the marginal product of labor equal to the nominal wage rate, $P_{i,t}^{L*}$ divided by $P_{i,t}^{Y*}$; this condition will be discussed in Section VI.C.) Differentiating equation (10) with respect to capital and rearranging terms (as detailed in Appendix I), we obtain the following factor demand equation for the long-run capital/output ratio,

$$(K_{i,t} / Y_{i,t})^* = \phi^\sigma \, ((P_{i,t}^K / P_{i,t}^Y)^*)^{-\sigma} \, U_{i,t}^{KY}, \qquad (11a)$$

$$U_{i,t}^{KY} \equiv A_{i,t}^{[\sigma-1]} \, B_t^{K[\sigma-1]}. \qquad (11b)$$

To capture fixed industry and aggregate effects, we assume that the error term follows a two-way error component model,

$$U_{i,t}^{KY} = \exp[u_i^{KY} + u_t^{KY} + u_{i,t}^{KY}], \qquad (12)$$

where $u_{i,t}^{KY}$ may have a non-zero mean. Taking logs of the first-order condition in equation (11a),

$$\ln(K_{i,t} / Y_{i,t})^* = \sigma \ln(\phi) - \sigma \ln(P_{i,t}^K / P_{i,t}^Y)^* + \ln U_{i,t}^{KY}, \qquad (13)$$

removing fixed industry effects by first-differencing, and defining
$ky_{i,t}^* \equiv \ln(K_{i,t} / Y_{i,t})^*$, $p_{i,t}^{KY*} \equiv \ln(P_{i,t}^K / P_{i,t}^Y)^*$, $\tau_t^{KY} \equiv \Delta u_t^{KY}$, and $e_{i,t}^{KY} \equiv \Delta u_{i,t}^{KY}$,
we obtain the following estimating equation,

$$\Delta ky_{i,t}^* = \zeta^{KY} - \sigma \Delta p_{i,t}^{KY*} + \tau_t^{KY} + e_{i,t}^{KY}, \qquad (14)$$

where τ_t^{KY} is an aggregate fixed time effect and ζ^{KY} is a constant term (included in place of one of the τ_t^{KY}'s). Conditional on time fixed effects, identification of σ is achieved by the correlation between the growth rates of the capital/output ratio and its relative price.

Given observations on $ky_{i,t}^*$, and $p_{i,t}^{KY*}$, equation (14) provides a rather straightforward framework for estimating σ. Consistent estimates are obtained because the industry-level factor prices are exogenous. (This assumption is relaxed in Section VI.A., which contains instrumental variables estimates.) Importantly, in light of the recent critique and evidence by Antrás (2004), our estimates of σ control for biased technical change.[7] The key unresolved issue is that the long-run values denoted by *'s are not observable, an issue to which we now turn.

---

[7] See his equation (1'), which is comparable to our equation (13) and, with first-differencing, equation (14). The effects of biased technical change are removed by a linear time trend in Antrás’ framework based on aggregate data and by time effects in our framework based on panel

III.B. Low-Pass Filters and Long-Run Values

Estimation of equation (14) is made difficult because the long-run values are not observed. Previous research has addressed this problem in several ways. The cointegration approach introduced by Caballero (1994, 1999) also focuses on long-run values and provides an elegant solution for extracting long-run values from data subject to short-run deviations. While innovative, this estimation strategy faces some econometric difficulties in recovering production function parameters (Chirinko and Mallick, 2007) and defines the long-run at only the 0^{th} frequency (as opposed to a broader band of frequencies; cf. fn. 10). Another approach estimates an investment equation that links changes in the observed capital stock to the unobserved long-run capital stock by assuming that 1) the change in the observed capital stock is measured by investment spending, 2) that the change in the long-run capital stock is measured by changes in output and the price of capital, and 3) that these changes are distributed over time due to various short-run frictions (Chirinko, 1993, Section II). Relying on investment data replaces the unobservability problem with a set of difficult issues concerning dynamics, transition toward a steady-state, and the specification of investment equations.[8]

Our approach also focuses on the first-order condition that holds in the long-run but uses the Baxter-King (1999) low-pass filter (LPF) to measure the long-run values of variables denoted by *'s. A LPF allows frequencies lower than some critical frequency, $\omega^{\#}$, to pass through to the transformed series but excludes

---

data. If we adopt Antrás' specification of biased technical change, $B_t^K = \exp[\lambda t]$, the λ is absorbed in the constant in equation (14). Panel data permit a more general specification to capture the effects of biased technical change.

[8] Chirinko, Fazzari, and Meyer (2004) document that, relative to the approach pursued in this paper, investment equations based on firm-level panel data impart a downward bias on estimates of σ.

frequencies higher than $\omega^{\#}$. Baxter and King present two important results regarding LPF's for the purpose of the current study. They derive the formulas that translate restrictions from the frequency domain into the time domain. For an input series, x_t, the ideal LPF for a critical value $\omega^{\#}$ produces the transformed series, $x_t^*[\omega^{\#}, q]$, for infinite lag and lead lengths, $q \rightarrow \infty$,

$$x_t^*[\omega^{\#}, q] = \lim_{q\rightarrow\infty} \sum_{h=-q}^{q} d_h[\omega^{\#}] x_{t-h}, \tag{15a}$$

$$d_h[\omega^{\#}] = d'_h[\omega^{\#}] + \theta[\omega^{\#}, q], \tag{15b}$$

$$d'_h[\omega^{\#}] = \omega^{\#} / \pi, \qquad h = 0, \tag{15c}$$

$$d'_h[\omega^{\#}] = \sin[|h|\,\omega^{\#}]/(|h|\,\pi), \qquad h = \pm 1, \pm 2, \ldots, q, \tag{15d}$$

$$\theta[\omega^{\#}, q] = \lim_{q\rightarrow\infty} \left(1 - \sum_{h=-q}^{q} d'_h[\omega^{\#}] \right) \Big/ (2q+1), \tag{15e}$$

$$\omega^{\#} = 2\pi / p^{\#} = F[p^{\#}], \qquad p^{\#} = [2, \infty), \tag{15f}$$

where the $d_h[\omega^{\#}]$'s are weights defined as the sum of two terms – a provisional set of weights denoted by a prime (the $d'_h[\omega^{\#}]$'s in equations (15c) and (15d)) and a frequently imposed normalization that the $d_h[\omega^{\#}]$'s sum to 1 (per the constant $\theta[\omega^{\#}, q]$ computed in equation (15e)). Equation (15g) defines the inverse relation between the critical frequency ($\omega^{\#}$) and the critical periodicity ($p^{\#}$), the latter defined as the length of time required for the series to repeat a complete cycle. Since periodicities are relatively easy to interpret, hereafter we focus on $p^{\#}$ in place of $\omega^{\#}$.

A difficulty with implementing equations (15) is that the *ideal* LPF requires an infinite amount of data. Baxter and King's second important result is that the *optimal approximate* LPF for a window (i.e., the length of the leads and lags) of finite length q truncates the symmetric moving average at q. Thus, for $|h| \le q$, the $d_h[p^\#]$'s are given in equations (15); for $|h| > q$, $d_h[p^\#] = 0$. The optimal approximate LPF for the critical periodicity $p^\#$ and lead and lag length q, $LPF[p^\#, q]$, is given by equations (15) for any finite q.

IV. Spectral Properties of the Low-Pass Filter

Our estimation strategy is designed to emphasize long-run variation, and this section uses spectral analysis to assess the extent to which choices of the critical periodicity and window are successful.[9] The estimating equation is derived in three steps: a) define long-run values with the $LPF[p^{\#}, q]$ (equations (15)); b) insert these long-run values into the first-order condition for optimal capital accumulation and take logarithms (equation (13)); c) first-difference this logarithmic equation to remove industry fixed effects (equation (14)). Each of these steps impacts the spectrum of the transformed data and hence the relative weights given to long-run variation in the variables ultimately entering the estimating equation. To compute the spectrum of a transformed series, we rely on the fundamental result from spectral analysis linking the spectrum of an output series to the product of the spectrum of an input series and a scalar that may be a function of ω, $p^{\#}$, or q. To understand the impact of each step, we need only compute the scalar associated with each transformation.

In analyzing the spectral properties of our estimator, it is convenient to recast the LPF transformation (for a finite q), the logarithmic transformation, and the first-difference transformation as follows,

$$x_t^*[p^{\#}, q] = \sum_{h=-q}^{q} d_h[p^{\#}] \; x_{t-h}, \tag{16a}$$

$$y_t^*[p^{\#}, q] = \ln[x_t^*[p^{\#}, q]], \tag{16b}$$

$$z_t^*[p^{\#}, q] = \Delta y_t^*[p^{\#}, q], \tag{16c}$$

---

[9] See Hamilton (1994, Chapter 6) or Sargent (1987, Chapter IX) for discussions of spectral analysis.

where x_t represents the raw data series, either $(K_{i,t} / Y_{i,t})$ or $p^K_{i,t}$. The spectra corresponding to the $x^*_t[.]$, $y^*_t[.]$, and $z^*_t[.]$ output series in equations (16) are defined over the interval ω = [0, π] as the product of the spectrum for the input series and a scalar that is nonnegative, real, and may be depend on ω, $p^\#$, or q,

$$g_{x*}[e^{-i\omega}] = a_\omega[p^\#, q]\ g_x[e^{-i\omega}], \tag{17a}$$

$$g_{y*}[e^{-i\omega}] = b\ g_{x*}[e^{-i\omega}], \tag{17b}$$

$$g_{z*}[e^{-i\omega}] = c_\omega\ g_{y*}[e^{-i\omega}], \tag{17c}$$

where $g_x[e^{-i\omega}]$ is the spectrum for the raw series and the scalars are defined as follows,

$$a_\omega[p^\#, q] = \alpha[p^\#, q] \left\{ \begin{array}{l} (2/p^\#) + 2\sum_{h=1}^{q} \cos[h\omega]\ d'_h[p^\#] \\ + \theta[p^\#, q]\ \{(1-\cos[\omega(2q+1)])/(1-\cos[\omega])\}^{1/2} \end{array} \right\}^2 \tag{17d}$$

$$b \ = \beta\ (\mu_{x*})^{-2}, \tag{17e}$$

$$c_\omega = \gamma\ 2\ (1-\cos[\omega]), \tag{17f}$$

where μ_{x*} equals the unconditional expectation of $x_t^*[.]$. To ensure comparability in the analyses to follow that vary $p^\#$ and q, the three spectra are normalized by an appropriate choice of a constant (α, β, or γ) so that the integrals for equations (17a), (17b), and (17c) evaluated from 0 to π equal 1.0.

The three scalars $a_\omega[p^\#,q]$, b, and c_ω correspond to the LPF, logarithmic, and first-difference transformations, respectively, and are derived as follows. The $a_\omega[p^\#,q]$ scalar is based on Sargent (1987, Chapter XI, equation (33)),

$$a_\omega[p^\#,q] = \alpha[p^\#,q] \left\{ \sum_{h=-q}^{q} e^{-ih\omega}\, d_h[p^\#] \right\} \left\{ \sum_{h=-q}^{q} e^{ih\omega}\, d_h[p^\#] \right\}. \tag{18}$$

The two-sided summations are symmetric about zero and only differ by the minus sign in the exponential terms. Hence, the two sums in braces are identical. The $d_h[.]$'s appearing in the summations are separated into $\theta[.]$ and the $d'_h[.]$'s (cf. equations (15)). For the latter terms, a further distinction is made between the term at h=0 and the remaining terms (h=±1,±q) that are symmetric about h=0. Equation (18) can be written as follows,

$$a_\omega[p^\#,q] = \alpha[p^\#,q] \left\{ \begin{array}{l} \theta[p^\#,q] \sum\limits_{h=-q}^{q} e^{ih\omega} \\ +(2/p^\#) + \sum\limits_{h=1}^{q} (e^{-ih\omega} + e^{ih\omega})\, d'_h[p^\#] \end{array} \right\}^2 . \tag{19}$$

The first sum of exponential terms is evaluated based on Sargent (1987, p. 275),

$$\sum_{h=-q}^{q} e^{ih\omega} = \left\{ \left(\sum_{h=-q}^{q} e^{ih\omega} \right)^2 \right\}^{1/2} \tag{20}$$
$$= \{((1-\cos[(2q+1)\omega]/(1-\cos[\omega])\}^{1/2}.$$

The second sum of exponential terms is evaluated with the Euler relations, $e^{\pm ih\omega} = \cos[h\omega] \pm \sin[h\omega]$,

$$\sum_{h=1}^{q} (e^{-ih\omega} + e^{ih\omega})\, d'_h[p^\#] = 2\sum_{h=1}^{q} \cos[h\omega]\; d'_h[p^\#]. \tag{21}$$

The b scalar is based on the approximation in Granger (1964, p. 48, equation 3.7.6), which states that the approximation will be accurate if the mean is much larger than the standard deviation of the input series ($x^*_t[.]$). The c_ω scalar is based on the well-known formula for the first-difference transformation (Hamilton, 1994, equation 6.4.8). Note that b and c_ω are independent of $p^\#$ and q. The importance of the above analytical results is that the combined effects of the three transformations are captured by three scalars that multiply the spectrum of the raw series,

$$g_{z*}[e^{-i\omega}] = \left\{ a_\omega[p^\#, q] * b * c_\omega \right\} * g_x[e^{-i\omega}]. \tag{22}$$

With equation (22), we are now in a position to examine the extent to which our estimation strategy based on the definition of the long-run ($p^\# = 8$) and a given window (q=3) is successful in emphasizing long-run variation in the data. Since the spectra for the raw series ($g_x[e^{-i\omega}]$) and the scalars associated with the logarithmic and first-difference transformations (b and c_ω, respectively) do not depend on $p^\#$ or q, they can be ignored in drawing relative comparisons among estimators. Alternative values of $p^\#$ or q will only affect the $LPF[p^\#, q]$ and the associated frequency response scalar, $a_\omega[p^\#, q]$.

Our first set of analyses holds the window fixed at q = 3 and examines different values of the critical periodicity, $p^\#$, that determines which frequencies are passed-through in the $LPF[p^\#, q]$. Four values of $p^\#$ are considered in Figure 2. We begin with the minimum value of the critical frequency, $p^\# = 2$, which corresponds to the standard investment equation that does not transform the raw data (other than the logarithmic and differencing operations). The frequency response for the standard investment model ($a_\omega[p^\# = 2, q = 3]$) is flat, indicating that this estimator does not reweight the variances across frequencies of the raw series. By contrast, our benchmark model represented by $a_\omega[p^\# = 8, q = 3]$ effects a substantial reweighting. With $p^\# = 8$, the benchmark model emphasizes the variances from periodicities greater than or equal to 8 years (which corresponds to $\omega^\# \leq 0.79$ on the horizontal axis), thus allocating a substantial amount of weight to those frequencies that we believe will yield better estimates of production function parameters. The remaining entries in Figure 2 are for the intermediate cases, $a_\omega[p^\# = 4, q = 3]$ and $a_\omega[p^\# = 6, q = 3]$.

The benchmark model is based on the assumption that periodicities greater than or equal to 8 years contain useful information for the parameter estimates. We now explore how much additional reweighting occurs when we increase the critical periodicity above 8; specifically, for values of $p^\#$ equal to 10, 20 and, in the limit, ∞. The $a_\omega[.]$'s corresponding to these critical values are graphed in Figure 3. Comparing the frequency responses for these higher periodicities indicates that they weight the lower frequencies in a manner very similar to the benchmark case of $p^{\#} = 8$.

This analysis suggests two conclusions concerning our choice of the critical periodicity. First, our estimation strategy based on $p^{\#} = 8$ appears to be reasonably successful in emphasizing long-run variation. This critical value appears to be a well-accepted standard for separating long-run frequencies from short-run and medium-run frequencies.[10] Second, results in Figure 3 suggest that parameter estimates are likely to be insensitive to the critical periodicity for values of $p^{\#} > 8$.

We can also use the spectral formulas to assess the impact of variations in the window, q, in approximating the ideal low-pass filter. Recall that the ideal LPF is based on the limiting behavior as q --> ∞. Since the span of our data are limited, this procedure is not feasible, and our empirical work relies on the optimal approximation based on a finite number of q leads and lags. This approximation introduces error into the analysis because variances associated with frequencies other than those desired enter into the transformation of the model variables. However, increasing q is costly in terms of lost degrees of freedom.

The tradeoff between approximation error and degrees of freedom is assessed in Figure 4, which plots $a_{\omega}[p^{\#} = 8, q]$ and values of q equal to 1, 3 and 5. When q = 1, the LPF is extensively contaminated by the variances associated with frequencies above the critical frequency of 0.79. As q increase to 3 and then 5, this contamination is reduced, and greater weight is placed on lower frequencies. Since using a window of q = 5 is costly in terms of degrees of freedom and roughly the

---

[10] A critical value of $p^{\#} = 8$ is used by Baxter and King (1999, p. 575), Levy and Dezhbakhsh (2003, p. 1502), Prescott (1986, p. 14), and Stock and Watson, 1999, p. 11). Burns and Mitchell (1946) report that the duration of the typical business cycle in the U.S. is less than 8 years.

same frequencies are emphasized, we will adopt q = 3 as our preferred window, though robustness will be examined with q = 1 and q = 5.[11]

[11] Baxter and King (1999, pp. 581-582) reached a similar conclusion based on their analysis of band-pass filters at middle frequencies.

V. Benchmark Empirical Results

This section estimates the substitution elasticity using our low-pass filter model defined with various critical periodicities and windows. Data are obtained from the webpage of Dale Jorgenson and represent output, inputs, and prices for 35 industries for the period 1959-1996.[12] The benchmark OLS results from our low-pass filtered model based on $p^{\#} = 8$ and $q = 3$ are as follows,

$$\Delta ky^{*}_{i,t} = \underset{(0.003)}{-0.003} - \underset{(0.020)}{0.308}\ \Delta p^{K*}_{i,t} + \tau^{K}_{t} + e^{K*}_{i,t} . \qquad R^2 = 0.404 \tag{23}$$

The point estimate for σ is 0.308 with a very small standard error of 0.020; the R^2 is 0.404. As we shall see, subsequent results very rarely depart in a meaningful way from the estimates presented in equation (23) and, when important differences occur, they are due to the presence of high frequency variation in the model variables.

Table 1 examines the sensitivity of estimates of σ to variations in the window (q) and the critical periodicity ($p^{\#}$). For a given $p^{\#}$, estimates of σ are very robust to variations in q. For example, when $p^{\#} = 8$, estimates of σ are 0.275, 0.308, and 0.292 for q of 1, 3, and 5, respectively. The standard errors rise as the window is increased and more data are used in computing the filters and less data are available for estimation. Nonetheless, the standard errors for the σ's remain less than 0.02 for all entries. (Note that the R^2's are not strictly comparable across cells because the dependent variable depends on $p^{\#}$ and q.) These results suggest

---

[12] The data are obtained at http://post.economics.harvard.edu/faculty/jorgenson/data/35klem.html and are described in Jorgenson and Stiroh (2000, especially the appendices). The effective time dimension equals the 38 datapoints contained in the dataset for a given industry less 2q for the construction of the LPF less one for first differencing.

that little is gained by increasing the size of the window and compromising degrees of freedom above q = 3.

Table 1 also allows us to assess the robustness to variations in $p^{\#}$ for a given q by reading down the columns.[13] For q = 3 in column 2, as $p^{\#}$ increase from 8 to ∞ in column 2, the estimates of σ hardly change. Consistent with the theoretical analysis in Figures 2 and 3, this robustness confirms that the relevant information about the long-run has been captured at $p^{\#} = 8$. When $p^{\#}$ is set to its minimum value of 2, the low-pass filter is neutral, and the raw data are transformed only by logarithmic and first-difference operations (cf. equation (22)). In this case, σ drops by one-third relative to the benchmark value (0.206 vs. 0.308). Thus, high frequency and presumably transitory variation affects the estimates of σ and, as has been frequently noted in the permanent income literature, transitory variation attenuates point estimates.

[13] Removal of the time dummies generally leads to a rise in σ (e.g., from 0.308 (Table 1, $p^{\#}$=8 and q=3) to 0.350) and a decline in the R^2 (from 0.404 to 0.335).

VI. Alternative Estimates

VI.A. Instrumental Variables

The estimates reported in Section V are based on the assumption that the long-run user cost of capital is unaffected by measurement error and hence that OLS is the appropriate estimation technique. Since this assumption necessary for consistent estimation may not hold strictly, instrumental variable estimates provide a useful robustness check.

Table 2 contains two sets of instrumental variable results. Panel A uses $p_{i,t-2}^{K*}$ as the instrument for the same range of values for $p^{\#}$ and q that appeared in Table 1. For values of $p^{\#}$ equal to or greater than 6, the IV estimates of σ are on average 0.050 higher than their OLS counterparts in Table 1 for the models using q = 3. For all models with $p^{\#}$ equal to or greater than 6, the average difference is 0.068. All of these estimated σ's are statistically significant at the 1% level. The gap widens for $p^{\#} = 4$. However, for $p^{\#} = 2$, the results become nonsensical. For those models emphasizing long-run variation, the IV estimates of σ are greater than the comparable OLS estimates, but still far from unity.

Recent work with instrumental variables has raised concerns about weak instruments and biased estimates (beginning with Nelson and Startz, 1990).[14] Instrumental relevance is assessed with the test statistic proposed by Stock, Wright, and Yogo (2002), which involves an auxiliary regression of the model variable on the instrument and a comparison of the F-statistic for the goodness of fit to a critical value of 8.96 (reported in their Table 1). As shown below, the σ's in Table 2, only the instrument used for the $p^{\#} = 2$ results is weak, and hence these estimates are unreliable. This result suggests a reason for the implausible estimates for $p^{\#}$ =

[14] Note that Hansen-Sargen test of instrument validity is not useful in this just identified model.

2. All of the other estimates are based on strong (relevant) instruments. For $p^{\#} \geq 4$, the estimated σ's range from 0.304 to 0.423.

OLS estimates may also be inconsistent if simultaneity induces a correlation between the error term and a well-measured regressor. In this case, $p_{i,t-2}^{K*}$ will not be a valid instrument because it contains leads that may also be correlated with the contemporaneous error term. While the model developed in Section III places restrictions on the construction of model variables, it does not place restrictions on the instrument, and thus we construct an alternative instrument, $p_{i,t-2}^{K\&}$, as a one-sided lag (see the note to Table 2 for details). In the face of simultaneity, this variable will be uncorrelated with the contemporaneous error term and will deliver consistent estimates.

Panel B of Table 2 repeats the models reported in Tables 1 and 2.A for a range of values of $p^{\#}$ and q. The σ's estimated with the $p_{i,t-2}^{K\&}$ instrument are quite similar to the other IV estimates in Table 2.A. The instrument relevance statistics continue to be large (relative to the critical value) for models based on $p^{\#} \geq 4$.

VI.B. Split-Samples

To further assess the robustness of our results, Table 3 contains OLS estimates from the first and second halves of the sample, 1959-1977 and 1978-1996, respectively. The results closely follow those reported previously. For example, for our preferred specification with $p^{\#} = 8$ and $q = 3$, the estimates of σ from the first and second halves of the sample are 0.261 and 0.362, respectively. These estimates are not economically different from our preferred estimate from the full sample of 0.308.

VI.C. Other Estimating Equations

The first-order conditions for profit maximization yield two additional estimating equations that contain the labor/output ratio or the labor/capital ratio as the dependent variable. However, the neoclassical growth model implies that neither series is stationary, an implication consistent with the first two of the stylized facts of growth advanced by Kaldor (1961) and analyzed statistically by Klein and Kosobud (1961). Hence, the low-pass filter used in this study is not strictly applicable because spectral methods require stationary data. Murray (2003) and Cogley and Nason (1995) document the problems that can arise when band-pass filters are applied to nonstationary data and Mallick (2006a) explores the effects of nonstationary data on the variety of estimates of σ appearing in the literature. With this important caveat noted, we nonetheless examine estimates of σ derived from the equations with labor/output, $\ell y^{*}_{i,t}$, or the capital/labor ratio, $k\ell^{*}_{i,t}$, as the dependent variable, and estimate the following equations,

$$\Delta \ell y^{*}_{i,t} = \zeta^{LY} - \sigma \Delta p^{LY*}_{i,t} + \tau^{LY*}_{t} + e^{LY*}_{i,t}, \qquad (24)$$

$$\Delta k\ell^{*}_{i,t} = \zeta^{KL} - \sigma \Delta p^{KL*}_{i,t} + \tau^{KL*}_{t} + e^{KL*}_{i,t}, \qquad (25)$$

where $p^{LY*}_{i,t} \equiv \ln(P^{L}_{i,t} / P^{Y}_{i,t})^{*}$, $p^{KL*}_{i,t} \equiv \ln(P^{K}_{i,t} / P^{L}_{i,t})^{*}$, and the other elements in equations (24) and (25) parallel those defined in equation (14).

Table 4 contains σ's and R^2's for q = 3 and the usual range of $p^{\#}$'s. Column 1 contains the previously reported estimates for $\Delta ky^{*}_{i,t}$ and columns 2 and 3 the results for $\Delta \ell y^{*}_{i,t}$ and $\Delta k\ell^{*}_{i,t}$, respectively. Relative to the results with $\Delta ky^{*}_{i,t}$, the σ's estimated with the $\Delta \ell y^{*}_{i,t}$ equation are higher for all critical periodicities (save

$p^{\#} = 4$); those for $\Delta k\ell^{*}_{i,t}$ are uniformly lower. The maximal difference among σ's across the three specifications is 0.17. The array of estimates of σ in Table 4 is bounded above by 0.40.

In a well-known study, Berndt (1976) estimated these three first-order conditions and uncovered a disturbingly wide range of results. Part of this dispersion was due to different definitions of factor prices. But he also found that the labor/output equation delivered higher values of σ. With the exception of the $p^{\#} = 4$ results, Table 4 confirms the Berndt finding.

VII. Heterogeneous Industry σ's

Given our interest in the impact of aggregate σ on growth theory, the homogeneity assumption imposed across industries in the prior two sections is a natural way of obtaining the substitution elasticity. From an estimation perspective, however, it might be desirable to exploit the panel feature of our dataset and to allow the σ's to differ across industries. Assuming that econometrically sound estimates of industry σ's (σ_i's) can be obtained, a problem arises because these σ_i's are not immediately useful in analyzing growth issues. A mapping is required from industry elasticities to the aggregate elasticity, σ_{agg}, relevant for growth theory. In the first sub-section, we use Hicks' formula for the derived demand of a factor of production to develop a mapping from the σ_i's to σ_{agg} that recognizes substitution within an industry and reallocations across industries.[15] The second sub-section presents estimates of the σ_i's and, based on the prior analytic results, the associated σ_{agg}.

VII.A. Mapping Between The Industry and Aggregate σ's

We begin by defining σ_{agg} for the representative aggregate firm as the percentage change in the ratio of capital and labor with respect to the percentage change in the ratio of the prices of capital and labor, holding output constant (Hicks, 1963, p. 289 and Robinson, 1933, p. 256). While it is quite tempting to compute σ_{agg} as a weighted-average (with weights ω_i) of the underlying industry σ_i's,

[15] We thank Robert Solow for highlighting the importance of reallocations in inferring an aggregate σ from disaggregate estimates and for posing the counter-example discussed below equation (26).

$$\sigma_{agg} = \sum_i \omega_i \sigma_i \,, \tag{26}$$

such a temptation should be resisted. Consider the case where the relative price of capital falls and all industries have Leontief technologies. Since $\sigma_i = 0$ for all industries, equation (26) implies that the aggregate substitution elasticity is zero. However, even with no substitution along their isoquants, industries would expand in the face of the lower factor price. The extent of this scale effect would depend on three industry characteristics: the importance of capital in the cost structure, the response of industry demand to a lower output price, and the relative size of the industry. Aggregate capital accumulation would increase initially. However, the definition given above requires that aggregate output be held constant. Imposing this constraint will lower capital accumulation, though only by happenstance will this reduction exactly extinguish the initial increase.

Thus, a mapping is required that clearly delineates, not only the effects of substitution within industries, but also the reallocations across industries, the latter defined as the difference between scale effects computed at the industry and aggregate levels. Our strategy is to start with an aggregate equation and then relate the aggregate and industry components. We begin with Hicks' formula for the derived demand of a factor of production applied to the representative aggregate firm and use this formulation to develop intuition and introduce definitions,

$$\lambda_{agg} = \sigma_{agg} - \sigma_{agg}\,\mu^{K}_{agg} + \eta_{agg}\,\mu^{K}_{agg}\,, \tag{27a}$$

$$\mu^{K}_{agg} \equiv \sum_{i} \omega_i\,\mu^{K}_{i}\,, \tag{27b}$$

$$\eta_{agg} \equiv \sum_{i} \omega_i\,\eta_i\,, \tag{27c}$$

where λ_{agg} is the aggregate own price elasticity of capital, μ^{K}_{agg} is the aggregate factor share for capital, η_{agg} is the aggregate price elasticity of output, and the ω_i's are weights.

The three terms in equation (27a) capture in a succinct manner the substitution and scale effects associated with a change in the price of capital for the representative aggregate firm. The first term captures the direct substitution effect holding output price and output constant. The second term represents an additional indirect substitution effect driven by the lower marginal cost of production. Under competitive conditions, the decline in marginal cost translates into a decline in the output price. The extent of this decline is determined by the relative importance of capital in production represented by μ^{K}_{agg}. Since output price enters the denominator of the user cost, the decline in output price raises the relative price of and lowers the demand for capital. The third effect occurs because the lower factor price raises demand. This scale effect is represented in the third term of equation (27a) by the product of μ^{K}_{agg} and the price elasticity of output, η_{agg}.

Equation (27a) can be solved for the aggregate substitution elasticity,

$$\sigma_{agg} = \left(\lambda_{agg} - \eta_{agg}\,\mu^{K}_{agg}\right)\Omega_{agg}\,, \tag{28a}$$

$$\Omega_{agg} \equiv \left(1 - \mu^{K}_{agg}\right)^{-1}. \tag{28b}$$

In equations (28), the aggregate substitution elasticity depends on the aggregate price elasticity less a subtraction for the aggregate scale effect, $\eta_{agg}\,\mu^{K}_{agg}$. As capital's share becomes vanishingly small ($\mu^{K}_{agg} \to 0$ and $\Omega_{agg} \to 1$) or demand becomes perfectly inelastic ($\eta_{agg} \to 0$), the scale effect disappears.

We relate industry parameters to equations (28) by beginning with a definition of the aggregate own price elasticity as a weighted-average of the industry own-price elasticity for a uniform percentage change in the price of capital (per Basu and Fernald, 1997, Section III),

$$\lambda_{agg} \equiv \frac{dK_{agg}/K_{agg}}{dP^{K}_{agg}/P^{K}_{agg}} = \sum_{i} \frac{dK_{i}/K_{agg}}{dP^{K}_{agg}/P^{K}_{agg}} = \sum_{i} (K_{i}/K_{agg}) \frac{dK_{i}/K_{i}}{dP^{K}_{agg}/P^{K}_{agg}} = \sum_{i} \omega_{i}\,\lambda_{i}\,, \tag{29a}$$

$$\omega_{i} \equiv K_{i}/K_{agg}\,, \tag{29b}$$

where the ω_i's are weights reflecting the size of industry i as measured by the percentage of total capital stock in industry i.[16] Applying Hicks' formula for industry i,

[16] An alternative aggregation scheme is developed by Houthakker (1956) and Jones (2005), who build a macro production function from micro production functions. Their approach has the benefit of developing solid microfoundations of production, but it tends to be more restrictive than the approach pursued here, which has the limited objective of mapping micro parameters into a macro parameter and identifying reallocation effects.

$$\lambda_i = \sigma_i - \sigma_i \mu_i^K + \eta_i \mu_i^K, \tag{30}$$

and substituting equation (30) into (29a) and the resulting expression into equation (28a), and rearranging terms, we obtain the following result,

RESULT I: *The mapping between the aggregate and industry substitution elasticities is given by the following formula that is the sum of substitution and reallocation effects,*

$$\sigma_{agg} = \underbrace{\left\{\sum_i \omega_i \sigma_i \left(1-\mu_i^K\right)\right\}\Omega_{agg}}_{Substitution} + \underbrace{\left\{\sum_i \omega_i \left(\eta_i \mu_i^K - \eta_{agg}\, \mu_{agg}^K\right)\right\}\Omega_{agg}}_{Reallocations}. \tag{31}$$

The substitution effect is represented by the first summation and equals the direct and indirect substitution effects at the industry level ($\sigma_i(1-\mu_i^K)$) weighted by industry size and adjusted by Ω_{agg}. The reallocation effect is represented by the second summation containing the difference in scale effects at the industry and aggregate levels (also adjusted by Ω_{agg}). This difference reflects the important characteristic that reallocations are measured by the industry scale effects relative to the aggregate scale effect.

Equation (31) captures two important effects. First, even if $\sigma_i = 0$ for all industries, σ_{agg} will not necessarily be zero because of reallocations. Second, in general, reallocations depend in a complicated manner on the interplay among the aggregate scale effect and the distributions of the industry weights (ω_i's) and industry scale effects ($\eta_i \mu_i^K$'s). The second term in equation (31) is not necessarily positive, and hence σ_{agg} can be smaller than the weighted-average of the σ_i 's.

Result I presents a general formula for mapping industry and aggregate elasticities. However, the presence of industry scale effects does not necessarily imply that reallocations affect the mapping from industry to aggregate substitution elasticities. Given plausible independence among the price elasticities, capital shares, and industry weights that reflect independence of supply technologies from demand preferences, reallocations do not affect this mapping.

RESULT II: *If (i) both the price elasticities of output at the industry level (η_i) and aggregate level (η_{agg})are independent of the weighted-capital share ($\omega_i \mu_i^K$) and (ii) the industry price elasticities are independent of the industry weights (ω_i), then the reallocation effect vanishes and the mapping from industry to aggregate parameters is given by the following equation,*

$$
\begin{aligned}
\sigma_{agg} &= \left\{ \sum_i \omega_i \, \sigma_i \left(1 - \mu_i^K \right) \right\} \Omega_{agg} \, , \\
&= \left\{ \sum_i \omega_i \, \sigma_i \frac{\left(1 - \mu_i^K \right)}{\left(1 - \mu_{agg}^K \right)} \right\} .
\end{aligned} \tag{32}
$$

The derivation of the unimportance of reallocations is contained in Appendix II. In light of Result II, the aggregate substitution elasticity is the weighted-average of the industry substitution elasticities adjusted by relative importance of the direct and indirect substitution effects.[17]

VII.B. Estimates of the σ_i's and the σ_{agg}

Given Result II, we are now in a position to relax the constraint that $\sigma_i = \sigma$ and map these industry estimates to σ_{agg}. Table 5 reports OLS results with the benchmark LPF parameters of $p^{\#} = 8$ and $q = 3$ for the heterogeneous σ_i's in columns 1 to 3. Entries in the first and second rows are weighted by capital and output weights, respectively. (While capital weights follow directly from the derivation, the results with output weights are presented as a robustness check.) Results for the homogeneous σ_i's (from Section V) are presented in columns 4 and 5. The estimated σ_{agg}'s from the homogeneous and heterogeneous models are quite similar in the preferred models with time effects (columns 1, 2, and 4). While the two aggregate estimates differ statistically, these differences are not economically meaningful.

There are two differences between the homogeneous and heterogeneous models. First, under homogeneity, the removal of time effects led to an increase in σ_{agg}. However, when the σ_i's vary across industries, the opposite result occurs. Second, as expected, standard errors rise for the heterogeneous estimates, and they are much larger for the heterogeneous model with time fixed effects. Part of this increase (relative to the homogeneous model) is driven by one large industry standard error (Finance, Insurance, and Real Estate). When this outlier is removed, the standard errors in column 1 are reduced by nearly 40% for the capital-weighted

---

[17] It should be noted that the σ_{agg} from the homogeneous model is also effectively a weighted estimate. The heterogeneous model weights the σ_i's by industry capital shares, while the homogeneous model weights the σ_i's by industry variances.

estimate and nearly 25% for the output-weighted estimate. The results in Table 5 confirm that σ_{agg} is far below the Cobb-Douglas value of unity.

7. Summary and Conclusions

The elasticity of substitution between labor and capital (σ) is a crucial parameter in growth theory. Solow's fundamental innovation can be cast in terms of σ, where the Harrod-Domar assumption that $\sigma = 0$ is replaced with the more general assumption that $\sigma \geq 0$. Our discussion highlights that a host of important growth issues depend on the precise positive value of σ. It affects the possibility of perpetual growth or decline, the level and growth of income per capital, the speed of convergence, the rate of return on capital, the role of biased technical change, and the relative roles of productive factors and technical efficiency in explaining differences in per capita income.

This crucial production function parameter is estimated by combining a low-pass filter with standard panel data techniques to identify the long-run relations appropriate to production function estimation. Our preferred point estimate is 0.30, and it proves robust to variations in several directions. Based on this estimate and our review of the growth literature, σ is not an engine of growth. This estimate is well below the critical value needed for perpetual growth in the neoclassical growth model. Moreover, the empirical results suggest that the dynamic macroeconomics in general and the growth literature in particular need to move away from the convenient but inaccurate assumption of σ equal to unity. Such a departure from a Cobb-Douglas production function will force an expansion of the neoclassical growth model to include, among other factors, a central role for biased technical change in influencing factor shares and balanced growth.

References

Acemoglu, Daron, “Labor- And Capital-Augmenting Technical Change,” *Journal of the European Economic Association* 1 (2003), 1-37.

Akerlof, George, and Nordhaus, William D., "Balanced Growth -- A Razor's Edge?," *International Economic Review* 8 (October 1967), 343-348.

Antrás, Pol, "Is the U.S. Aggregate Production Function Cobb-Douglas?: New Estimates of the Elasticity of Substitution," *Contributions to Macroeconomics* 4 (2004), Article 4.

Barro, Robert J., and Sala-i-Martin, Xavier, *Economic Growth* (New York: McGraw-Hill, 1995).

Basu, Susanto, and Fernald, John G., "Returns to Scale in U.S. Production: Estimates and Implications," *Journal of Political Economy* 105 (April 1997), 249-283.

Baxter, Marianne, and King, Robert G., "Measuring Business Cycles: Approximate Band-Pass Filters for Economic Time Series," *The Review of Economics and Statistics* 81 (November 1999), 575-593.

Berndt, Ernst R., "Reconciling Alternative Estimates of the Elasticity of Substitution," *The Review of Economics and Statistics* 63 (February 1976), 59-68.

Blackorby, Charles, and Russell, R. Robert, “Will the Elasticity of Substitution Please Stand Up? (A Comparison of the Allen/Uzawa and Morishima Elasticities),” *American Economic Review* 79 (September 1989), 882-888.

Brown, Murray, *On the Theory and Measurement of Technological Change* (Cambridge: Cambridge University Press, 1968).

Burmeister, Edwin and Dobell, A. Rodney, *Mathematical Theories of Economic Growth*, (London: The Macmillian Company/Collier-Macmillan Limited, 1970).

Burns, Arthur M. and Mitchell, Wesley C., *Measuring Business Cycles* (New York: NBER, 1946).

Caballero, Ricardo J., "Small Sample Bias And Adjustment Costs," *The Review Of Economics And Statistics* 76 (February 1994), 52-58.

Caballero, Ricardo J., "Aggregate Investment," in John B. Taylor and Michael Woodford (eds.), *Handbook Of Macroeconomics,* Volume 1B (Amsterdam: Elsevier North-Holland), 1999), 813-862.

Caselli, Francesco, "Accounting for Cross-Country Income Differences," in Philippe Aghion and Steven Durlauf (eds.), *Handbook of Economic Growth*, Volume 1A (Amsterdam: Elsevier North-Holland, 2005), 679-741.

Caselli, Francesco, and Feyrer, "The Marginal Product of Capital," *Quarterly Journal of Economics* (forthcoming).

Chirinko, Robert S., "Business Fixed Investment Spending: Modeling Strategies, Empirical Results, and Policy Implications," *Journal of Economic Literature* 31 (December 1993), 1875-1911.

Chirinko, Robert S., "Corporate Taxation, Capital Formation, and the Substitution Elasticity between Labor and Capital," *National Tax Journal* 55 (June 2002), 339-355.

Chirinko, Robert S., Fazzari, Steven M., and Meyer, Andrew P., "That Elusive Elasticity: A Long-Panel Approach To Estimating The Capital-Labor Substitution Elasticity," Emory University (January 2004).

Chirinko, Robert S., and Mallick, Debdulal, "The Fisher/Cobb-Douglas Paradox, Factor Shares, And Cointegration," Emory University (January 2007).

Cogley, Timothy, and Nason, Jason M., "Effects of the Hodrick-Prescott Filter on Trend and Difference Stationary Time Series: Implications for Business Cycle Research," *Journal of Economic Dynamics and Control* 19 (January/ February 1995), 253-278.

Eisner, Robert, "A Permanent Income Theory for Investment: Some Empirical Explorations," *American Economic Review* 57 (June 1967), 363-390.

Engle, Robert F., and Foley, Duncan K., "An Asset Price Model of Aggregate Investment," *International Economic Review* 16 (December 1975), 625-647.

Friedman, Milton, *A Theory of the Consumption Function* (Princeton: Princeton University Press (for the NBER), 1957).

Granger, Clive W.J. (in association with Michio Hatanaka), Spectral Analysis of Economic Time Series (Princeton: Princeton University Press, 1964).

Hamilton, James D., *Time Series Analysis* (Princeton: Princeton University Press, 1994).

Hicks, John R., *The Theory of Wages*, Second Edition (London: MacMillan & Co., 1963). First edition published in 1932.

Houthakker, Hendrik S., "The Pareto Distribution and the Cobb-Douglas Production Function in Activity Analysis," *Review of Economic Studies* 23 (1955/1956), 27-31.

Jones, Charles I., "The Shape of Production Functions and the Direction of Technical Change," *Quarterly Journal of Economics* 120 (May 2005), 517-550.

Jones, Ronald W., and Ruffin, Roy J., “Trade and Wages: A Deeper Investigation,” University of Houston and University of Rochester (October 2003).

Jorgenson, Dale W., and Stiroh, Kevin J., "Raising the Speed Limit: U.S. Economic Growth in the Information Age," *Brookings Papers on Economic Activity* (2000:1), 125-211.

Kaldor, Nicholas, "Capital Accumulation and Economic Growth," in Freidrich A. Lutz and Douglas C. Hague (eds.), *The Theory of Capital* (New York: St. Martin's Press, 1961), 177-222.

King, Robert G., and Rebelo, Sergio T., "Transitional Dynamics and Economic Growth in the Neoclassical Model," *American Economic Review* 83 (September 1993), 908-932.

King, Robert G., and Rebelo, Sergio T., “Resuscitating Real Business Cycles,” in John B. Taylor and Michael Woodford (eds.), *Handbook Of Macroeconomics,* Volume 1B (Amsterdam: Elsevier North-Holland), 1999), 927-1007.

Klein, Lawrence, R., and Kosobud, Richard F., "Some Econometrics of Growth: Great Ratios of Economics," *Quarterly Journal of Economics* 75 (May 1961), 173-198. Reprinted in Jaime Marquez (ed.), *Economic Theory and Econometrics* (Philadelphia: University of Pennsylvania Press, 1985), 288-313.

Klump, Rainer, and de La Grandville, Oliver, “Economic Growth and the Elasticity of Substitution: Two Theorems and Some Suggestions,” *American Economic Review* 90 (March 2000), 282-291.

Klump, Rainer, and Preissler, Harald, “CES Production Functions and Economic Growth,” *Scandinavian Journal of Economics* 102 (March 2000), 41-56.

de La Grandville, Oliver, “In Quest of the Slutsky Diamond,” *American Economic Review* 79 (June 1989), 468-481.

de La Grandville, Oliver, and Solow, Robert M., “On the Determinants of Economic Growth: Is Something Missing?,” University of Geneva and MIT (April 2004).

Levy, Daniel, and Dezhbakhsh, Hashem, "International Evidence on Output Fluctuation and Shock Persistence," *Journal of Monetary Economics* 50 (October 2003), 1499-1530.

Lucas, Robert E., Jr., "Why Doesn't Capital Flow from Rich to Poor Countries?," *American Economic Review* 80 (May 1990), 92-96.

Mallick, Debdulal, "Reconciling Different Estimates of the Elasticity of Substitution," Emory University (June 2006a).

Mallick, Debdulal, "What Role for the Elasticity of Substitution in Economic Growth?: A Cross-Country Test of the Grandville Hypothesis," Emory University (July 2006b).

Mankiw, N. Gregory, “The Growth Of Nations,” *Brookings Papers On Economic Activity* (1995:1), 275-310.

Miyagiwa, Kaz and Papageorgiou, Chris, “Elasticity Of Substitution And Growth: Normalized CES In The Diamond Model,” *Economic Theory* 21 (January 2003), 155-165.

Murray, Christian J., "Cyclical Properties of Baxter-King Filtered Time Series," *Review of Economics and Statistics* 85 (May 2003), 472-476.

Nelson, Charles R., and Startz, Richard, "Some Further Results on the Exact Small Sample Properties of the Instrumental Variables Estimator," *Econometrica* 58 (July 1990), 967-976.

Pitchford, John D., "Growth And The Elasticity Of Substitution," *Economic Record* 36 (December 1960), 491-504.

Prescott, Edward C., "Theory Ahead of Business-Cycle Measurement," in Karl Brunner and Allan H. Meltzer (eds.), *Real Business Cycles, Real Exchange Rates and Actual Policies*, Carnegie-Rochester Series on Public Policy 25 (Autumn 1986), 11-44.

Ramanathan, R., "The Elasticity of Substitution and the Speed of Convergence in Growth Models," *Economic Journal* 85 (September 1975), 612-613.

Robinson, Joan, *The Economics of Imperfect Competition* (London: MacMillan & Co., 1933; Reprinted 1959).

Sargent, Thomas J., *Macroeconomic Theory* Second Edition (Boston: Academic Press, 1987).

Solow, Robert M., "A Contribution to the Theory of Economic Growth," *Quarterly Journal of Economics* 70 (1956), 65-94.

Stock, James H., and Watson, Mark W., "Business Cycle Fluctuations in US Macroeconomic Time Series," in John B. Taylor and Michael Woodford (eds.), *Handbook Of Macroeconomics,* Volume 1A (Amsterdam: Elsevier North-Holland), 1999), 3-64.

Stock, James H., Wright, Jonathan H., and Yogo, Motohiro, "A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments," *Journal of Business & Economic Statistics* 20 (October 2002), 518-529.

Turnovsky, Stephen J, "Intertemporal and Intratemporal Substitution, and the Speed of Convergence in the Neoclassical Growth Model," *Journal of Economic Dynamics & Control* 26 (August 2002), 1765-1785.

White, Halbert, "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity," *Econometrica* 48 (May 1980), 817-838.

White, Halbert, "Instrumental Variables Regression with Independent Observations," *Econometrica* 50 (March 1982), 483-500.

Yuhn, Ku-huang, “Economic Growth, Technical Change Biases, and the Elasticity of Substitution: A Test of the De La Grandville Hypothesis,” *Review of Economics and Statistics* 73 (May 1991), 340-346.

Appendix I:

Specifying the Marginal Product of Capital

With Neutral and Biased Technical Change

This appendix presents the details of the derivation of the marginal product of capital when there is both neutral and biased technical change. We assume that production possibilities are described by the following CES technology that relates output ($Y^*_{i,t}$) to capital ($K^*_{i,t}$), labor ($L^*_{i,t}$), neutral technical progress ($A_{i,t}$), and biased technical progress on capital and labor (B^K_t and B^L_t, respectively) for firm i at time t,

$$
\begin{aligned}
Y^*_{i,t} &= Y[K^*_{i,t}, L^*_{i,t}, A_{i,t}, B^K_t, B^L_t], \qquad \text{(A1)}\\
&= A_{i,t}\left\{\phi(B^K_t K^*_{i,t})^{[(\sigma-1)/\sigma]} + (1-\phi)(B^L_t L^*_{i,t})^{[(\sigma-1)/\sigma]}\right\}^{[\sigma/(\sigma-1)]}
\end{aligned}
$$

where ϕ is the capital distribution parameter and σ is the elasticity of substitution between labor and capital.

The derivative of $Y^*_{i,t}$ with respect to $K^*_{i,t}$ is computed from equation (A1) as follows,

$$
\begin{aligned}
Y'^*_{i,t} = {} & [\sigma/(\sigma-1)]A_{i,t}\left\{\phi(B^K_t K^*_{i,t})^{[(\sigma-1)/\sigma]} + (1-\phi)(B^L_t L^*_{i,t})^{[(\sigma-1)/\sigma]}\right\}^{[[\sigma/(\sigma-1)]-1]} \\
& *[(\sigma-1)/\sigma]\phi(B^K_t K^*_{i,t})^{[[(\sigma-1)/\sigma]-1]}B^K_t. \qquad \text{(A2)}
\end{aligned}
$$

Noting that the set of parameters in the exponent of $(B^K_t K^*_{i,t})$ on the second line of equation (A2) can be rewritten,

$$((\sigma-1)/\sigma)-1=-1/\sigma, \tag{A3}$$

we rearrange equation (A2) as follows,

$$\begin{aligned} Y'^{*}_{i,t} &= \phi\, K^{*\,[-1/\sigma]}_{i,t} \\ &\quad A_{i,t}\left\{\phi(B^{K}_{t}K^{*}_{i,t})^{[(\sigma-1)/\sigma]}+(1-\phi)(B^{L}_{t}L^{*}_{i,t})^{[(\sigma-1)/\sigma]}\right\}^{[\sigma/(\sigma-1)]} \\ &\quad \left\{\phi(B^{K}_{t}K^{*}_{i,t})^{[(\sigma-1)/\sigma]}+(1-\phi)(B^{L}_{t}L^{*}_{i,t})^{[(\sigma-1)/\sigma]}\right\}^{-1} \\ &\quad B^{K[(\sigma-1)/\sigma]}_{t}. \end{aligned} \tag{A4}$$

In equation (A4), the second line equals $Y^{*}_{i,t}$ per equation (A1), and the third line equals the product of $Y^{*}_{i,t}$ and $A_{i,t}$ raised to the appropriate powers,

$$\begin{aligned} Y'^{*}_{i,t} &= \phi\, K^{*\,[-1/\sigma]}_{i,t} \\ &\quad Y^{*}_{i,t} \\ &\quad Y^{*[(1-\sigma)/\sigma]}_{i,t}\; A^{[(\sigma-1)/\sigma]}_{i,t} \\ &\quad B^{K[(\sigma-1)/\sigma]}_{t}, \end{aligned} \tag{A5}$$

which can be rewritten as follows,

$$Y'^{*}_{i,t} = \phi\, K^{*\,[-1/\sigma]}_{i,t}\, Y^{*[1/\sigma]}_{i,t}\, A^{[(\sigma-1)/\sigma]}_{i,t}\; B^{K\,[(\sigma-1)/\sigma]}_{t} \quad . \tag{A6a}$$

$$= \phi\, ((Y_{i,t}/K_{i,t})^{*})^{[1/\sigma]}\, U^{KY\,[1/\sigma]}_{i,t}, \tag{A6b}$$

$$U^{KY\,[1/\sigma]}_{i,t} \equiv A^{[\sigma-1]}_{i,t}\; B^{K[\sigma-1]}_{t}. \tag{A6c}$$

We assume that the marginal product of capital in equation (A6a) is equated to the user cost of capital, the price of capital divided by the price of output,

$$(P^{K}_{i,t} / P^{Y}_{i,t})^{*} = \phi\, ((Y_{i,t} / K_{i,t})^{*})^{[1/\sigma]}\ U^{KY\,[1/\sigma]}_{i,t}. \tag{A7}$$

Equation (A7) can be rearranged to isolate the capital/output ratio on the left-side,

$$(K_{i,t} / Y_{i,t})^{*} = \phi^{\sigma}\ ((P^{K}_{i,t} / P^{Y}_{i,t})^{*})^{-\sigma}\ U^{KY}_{i,t}\,, \tag{A8}$$

which is equation (11a) in the text.

Appendix II:
The Unimportance of Reallocations

This Appendix demonstrates that the expected value of the reallocation term in equation (31) vanishes under the following two independence assumptions that reflect independence of supply technologies from demand preferences:

Assumption (i): Both the price elasticities of output at the industry (η_i) and aggregate (η_{agg}) levels are independent of the weighted-capital share ($\omega_i \mu_i^K$).

Assumption (ii): The industry price elasticities are independent of the industry weights (ω_i),

We begin by taking the expected value of the reallocation term in equation (31),

$$E\left\{\sum_i \omega_i\left(\eta_i\mu_i^K - \eta_{agg}\,\mu_{agg}^K\right)\Omega_{agg}\right\} = \tag{B1}$$

$$E\left\{\sum_i\left(\omega_i\eta_i\mu_i^K\right) - \eta_{agg}\,\mu_{agg}^K\right\}\Omega_{agg}\,.$$

Adding and subtracting η_{agg} to equation (B1) and rearranging, we obtain the following equation,

$$E\left\{\sum_i\left(\omega_i\mu_i^K\left(\eta_i - \eta_{agg}\right)\right) + \eta_{agg}\sum_i\left(\omega_i\mu_i^K\right) - \eta_{agg}\,\mu_{agg}^K\right\}\Omega_{agg} = \tag{B2}$$

$$E\left\{\sum_i \omega_i \mu_i^K \left(\eta_i - \eta_{agg}\right)\right\}\Omega_{agg}.$$

Given assumption (i), equation (B2) can be written as follows,

$$\left\{\sum_i E\left\{\omega_i \mu_i^K\right\}\left(E\{\eta_i\} - E\left\{\eta_{agg}\right\}\right)\right\}\Omega_{agg}. \tag{B3}$$

Given assumption (ii) and the constraint that the weights sum to unity, the expectation of η_{agg} can be written as follows,

$$E\left\{\eta_{agg}\right\} = \sum_{i=1}^{I} E\{\omega_i \eta_i\} = \sum_{i=1}^{I-1} E\left\{\omega_i\left(\eta_i - \eta_I\right)\right\} + E\left\{\eta_I\right\} = E\{\eta\}. \tag{B4}$$

Consequently, equation (B3) is zero, the expected value of the reallocation term in equation (B1) vanishes, and Result II follows immediately from Result I in the text.

Table 1: Ordinary Least Squares Estimates Of Equation (14) Dependent Variable: Capital/Output Ratio Various Critical Periodicities ($p^{\#}$) and Windows (q)

		q = 1	q = 3	q = 5
		(1)	(2)	(3)
$p^{\#} = 2$	σ_{agg} {R^2}	0.206 {0.466}	0.206 {0.476}	0.218 {0.495}
$p^{\#} = 4$	σ_{agg} {R^2}	0.275 {0.500}	0.278 {0.519}	0.281 {0.532}
$p^{\#} = 6$	σ_{agg} {R^2}	0.276 {0.487}	0.296 {0.454}	0.274 {0.402}
$p^{\#} = 8$	σ_{agg} {R^2}	0.275 {0.483}	**0.308 {0.404}**	0.292 {0.381}
$p^{\#} = 10$	σ_{agg} {R^2}	0.274 {0.481}	0.311 {0.390}	0.314 {0.359}
$p^{\#} = 20$	σ_{agg} {R^2}	0.274 {0.480}	0.304 {0.392}	0.327 {0.349}
$p^{\#} \to \infty$	σ_{agg} {R^2}	0.273 {0.480}	0.302 {0.394}	0.313 {0.353}

Notes: Estimates of σ_{agg} are based on panel data for 35 industries for the period 1959-1996. The effective time dimension equals the 38 datapoints contained in the dataset for a given industry less 2q for the construction of the LPF less one for first differencing. Standard errors are heteroscedastic consistent using the technique of White (1980), are less than or equal to 0.02 for all entries, and are not reported because all estimates of σ_{agg} are statistically significant at the 1% level. A constant term and fixed time effects (the number of fixed time effects equals the effective time dimension less one due to the inclusion of the constant) are included in the regression equation but are not reported. The R^2's are not comparable across cells because the dependent variable depends on $p^{\#}$ and q. Our preferred estimate is for the equation for which the Low-Pass Filter parameters are $p^{\#} = 8$ and q = 3.

Table 2: Instrumental Variable Estimates Of Equation (14) Dependent Variable: Capital/Output Ratio Various Critical Periodicities ($p^{\#}$) and Windows (q)

A. Instrument: $p^{KY*}_{i,t-2}$

		q = 1	q = 3	Q = 5
		(1)	(2)	(3)
$p^{\#} = 2$	σ_{agg} [F]	-0.575$^{@}$ [6.88]	-0.605$^{@}$ [6.82]	1.560$^{@}$ [7.48]
$p^{\#} = 4$	σ_{agg} [F]	0.388 [25.33]	0.385 [25.73]	0.423 [21.00]
$p^{\#} = 6$	σ_{agg} [F]	0.389 [27.00]	0.342 [58.16]	0.304 [58.97]
$p^{\#} = 8$	σ_{agg} [F]	0.393 [25.90]	**0.331 [118.56]**	0.328 [74.97]
$p^{\#} = 10$	σ_{agg} [F]	0.396 [25.35]	0.340 [117.19]	0.338 [147.91]
$p^{\#} = 20$	σ_{agg} [F]	0.398 [24.78]	0.374 [57.04]	0.352 [129.11]
$p^{\#} \rightarrow \infty$	σ_{agg} [F]	0.399 [24.69]	0.382 [48.30]	0.362 [70.47]

B. Instrument: $p^{KY\&}_{i,t-2}$

		q = 1	q = 3	Q = 5
		(1)	(2)	(3)
$p^{\#} = 2$	σ_{agg} [F]	-0.575@ [6.43]	-0.606@ [6.48]	0.827 [7.35]
$p^{\#} = 4$	σ_{agg} [F]	0.396 [23.82]	0.383[26.54]	0.419 [21.80]
$p^{\#} = 6$	σ_{agg} [F]	0.375 [29.96]	0.342 [56.24]	0.312 [55.24]
$p^{\#} = 8$	σ_{agg} [F]	0.373 [30.81]	**0.313 [75.73]**	0.328 [71.97]
$p^{\#} = 10$	σ_{agg} [F]	0.372 [31.00]	0.337 [85.33]	0.336 [92.24]
$p^{\#} = 20$	σ_{agg} [F]	0.372 [31.10]	0.357 [52.58]	0.361 [73.37]
$p^{\#} \rightarrow \infty$	σ_{agg} [F]	0.397 [25.11]	0.388 [54.65]	0.372 [65.36]

Notes: Estimates of σ_{agg} are based on panel data for 35 industries for the period 1959-1996. The effective time dimension equals the 38 datapoints contained in the dataset for a given industry less 2q for the construction of the LPF less one for first differencing less one for the instrument lagged two periods. The instrument is $p^{KY*}{}_{i,t-2}$ in panel A and $p^{KY\&}{}_{i,t-2}$ in panel B. The $p^{KY\&}{}_{i,t-2}$ instrument is constructed from equation (15a) with h = {0, −q}. Standard errors are heteroscedastic consistent using the technique of White (1982) and are not reported because all estimates of σ_{agg} are statistically significant at the 1% level with the exception of the σ_{agg}'s in row 1 marked with a $^{@}$, which are not significant at the 10% level. [F] is the F-statistic for the first-stage regression of $\Delta p^{KY*}{}_{i,t}$ on $p^{KY*}{}_{i,t-2}$ in panel A and $\Delta p^{KY*}{}_{i,t}$ on $p^{KY\&}{}_{i,t-2}$ in panel B. The null hypothesis of a weak instrument is rejected at the 5% level for F greater than or equal to 8.96 (Stock, Wright, and Yogo, 2002, Table 1). A constant term and fixed time effects (the number of fixed time effects equals the effective time dimension less one due to the inclusion of the constant) are included in the regression equation but are not reported. Our preferred estimate is for the equation for which the Low-Pass Filter parameters are $p^{\#} = 8$ and $q = 3$.

Table 3: Ordinary Least Squares Estimates Of Equation (14) Dependent Variable: Capital/Output Ratio Various Critical Periodicities ($p^{\#}$) and q = 3 Split-Sample Results

	Period		q = 1	q = 3	q = 5
			(1)	(2)	(3)
$p^{\#} = 2$	1959-77	σ_{agg} {R^2}	0.199 {0.536}	0.208 {0.517}	0.236 {0.527}
	1978-96	σ_{agg} {R^2}	0.203 {0.413}	0.179 {0.401}	0.118 {0.364}
$p^{\#} = 4$	1959-77	σ_{agg} {R^2}	0.243 {0.556}	0.250 {0.530}	0.270 {0.505}
	1978-96	σ_{agg} {R^2}	0.291 {0.471}	0.268 {0.386}	0.216 {0.298}
$p^{\#} = 6$	1959-77	σ_{agg} {R^2}	0.236 {0.540}	0.248 {0.436}	0.294 {0.434}
	1978-96	σ_{agg} {R^2}	0.301 {0.472}	0.326 {0.367}	0.292 {0.321}
$p^{\#} = 8$	**1959-77**	σ_{agg} {R^2}	0.232 {0.533}	**0.261 {0.410}**	0.315 {0.451}
	1978-96	σ_{agg} {R^2}	0.302 {0.470}	**0.362 {0.357}**	0.300 {0.283}
$p^{\#} = 10$	1959-77	σ_{agg} {R^2}	0.231 {0.531}	0.273 {0.423}	0.339 {0.415}
	1978-96	σ_{agg} {R^2}	0.302 {0.470}	0.367 {0.349}	0.355@ {0.241}
$p^{\#} = 20$	1959-77	σ_{agg} {R^2}	0.230 {0.528}	0.285 {0.458}	0.336 {0.399}
	1978-96	σ_{agg} {R^2}	0.302 {0.469}	0.344 {0.337}	0.394@ {0.272}
$p^{\#} \rightarrow \infty$	1959-77	σ_{agg} {R^2}	0.229 {0.528}	0.286 {0.465}	0.317 {0.403}
	1978-96	σ_{agg} {R^2}	0.302 {0.469}	0.337 {0.335}	0.379@ {0.298}

Notes: Estimates of σ_{agg} are based on panel data for 35 industries for the sub-samples indicated in the row headings. The effective time dimension equals the 19 datapoints contained in the dataset for each sub-sample for a given industry less 2q for the construction of the LPF less one for first differencing. The data are filtered after the sub-sample is defined. Standard errors are heteroscedastic consistent using the technique of White (1980), are less than or equal to 0.02 for all entries with the exception of the σ_{agg}'s in column 3 marked with a @), and are not reported because all estimates of σ_{agg} are statistically significant at the 1% level. A constant term and fixed time effects (the number of fixed time effects equals the effective time dimension less one due to the inclusion of the constant) are included in the regression equation but are not reported. The R^2's are not comparable across cells because the dependent variable depends on $p^{\#}$ and q. Our preferred estimate is for the equation for which the Low-Pass Filter parameters are $p^{\#} = 8$ and q = 3.

Table 4: Ordinary Least Squares Estimates Of Equations (14), (24), and (25) Alternative Dependent Variables Various Critical Periodicities ($p^{\#}$) and q = 3

		Dependent Variable		
		$\Delta ky^*_{i,t}$	$\Delta \ell y^*_{i,t}$	$\Delta k\ell^*_{i,t}$
		(1)	(2)	(3)
$p^{\#} = 2$	σ_{agg} $\{R^2\}$	.206 {.476}	.233 {.214}	.109 {.417}
$p^{\#} = 4$	σ_{agg} $\{R^2\}$	.278 {.519}	.239 {.249}	.168 {.419}
$p^{\#} = 6$	σ_{agg} $\{R^2\}$	.296 {.454}	.298 {.276}	.207 {.393}
$p^{\#} = 8$	σ_{agg} $\{R^2\}$	**.308 {.404}**	.353 {.310}	.245 {.387}
$p^{\#} = 10$	σ_{agg} $\{R^2\}$	.311 {.390}	.379 {.328}	.255 {.384}
$p^{\#} = 20$	σ_{agg} $\{R^2\}$	.304 {.392}	.398 {.341}	.239 {.377}
$p^{\#} \rightarrow \infty$	σ_{agg} $\{R^2\}$	.302 {.394}	.399 {.340}	.233 {.375}

Notes: Estimates of σ_{agg} are based on panel data for 35 industries for the period 1959-1996. The effective time dimension equals the 38 datapoints contained in the dataset for a given industry less 2q for the construction of the LPF less one for first differencing. Standard errors are heteroscedastic consistent using the technique of White (1980), are less than or equal to 0.02 for all entries, and are not reported because all estimates of σ_{agg} are statistically significant at the 1% level. A constant term and fixed time effects (the number of fixed time effects equals the effective time dimension less one due to the inclusion of the constant) are included in the regression equation, but are not reported. The R^2's are not comparable across cells because the dependent variable depends on $p^{\#}$ and q. The window is fixed at q = 3; our preferred estimate is for the equation in column (1) for which the Low-Pass Filter parameter is $p^{\#} = 8$.

Table 5: Ordinary Least Squares Estimates Of Equation (14) Dependent Variable: Capital/Output Ratio $p^{\#}$=8 and q=3 With And Without Fixed Time Effects Alternative Industry Weights

<table>
<tr><th></th><th colspan="3">Heterogeneous</th><th colspan="2">Homogeneous</th></tr>
<tr><td>Weights</td><td>With Time Effects</td><td>With Time Effects (One Outlier Removed)</td><td>Without Time Effects</td><td>With Time Effects</td><td>Without Time Effects</td></tr>
<tr><td></td><td>(1)</td><td>(2)</td><td>(3)</td><td>(4)</td><td>(5)</td></tr>
<tr><td></td><td></td><td></td><td></td><td rowspan="6">0.308
(0.020)</td><td rowspan="6">0.350
(0.018)</td></tr>
<tr><td>Capital</td><td>0.368</td><td>0.368</td><td>0.234</td></tr>
<tr><td></td><td>(0.183)</td><td>(0.112)</td><td>(0.024)</td></tr>
<tr><td></td><td></td><td></td><td></td></tr>
<tr><td>Output</td><td>0.365</td><td>0.365</td><td>0.271</td></tr>
<tr><td></td><td>(0.111)</td><td>(0.085)</td><td>(0.021)</td></tr>
</table>

Notes: Estimates of σ_{agg} are based on panel data for 35 industries for the period 1959-1996. The effective time dimension equals the 38 datapoints contained in the dataset for a given industry less 2q (=6) for the construction of the LPF less one for first differencing. Standard errors are in parentheses and are computed with the following formula, $\sqrt{\sum_i \omega_i^2 \, \mathrm{VAR}[\sigma_i]}$, where the $\mathrm{VAR}[\sigma_i]$'s are heteroscedastic consistent using the technique of White (1980). The results in column 2 exclude the VAR[$\sigma_{\text{Finance,Insurance,Real Estate}}$]. The weights, ω_i's, are defined as follows, $\omega_i \equiv \sum_t X_{i,t} / \sum_t \sum_i X_{i,t}$, where $X_{i,t} = K_{i,t}$ and $X_{i,t} = Y_{i,t}$ for the capital and output weights, respectively. A constant term and fixed time effects (the number of fixed time effects equals the effective time dimension less one due to the inclusion of the constant) are included in the regressions equations reported in columns 1 and 3; only a constant term is included in the regressions in columns 2 and 4. The Low-Pass Filter parameters are $p^{\#}$ = 8 and q = 3.

Figure 1
Steady-State Relation Between

Growth in the Capital/Labor Ratio $\left(\dot{(K/L)}/(K/L)\right)$ and σ
Equations (1), (5), (6), (7), and (8)

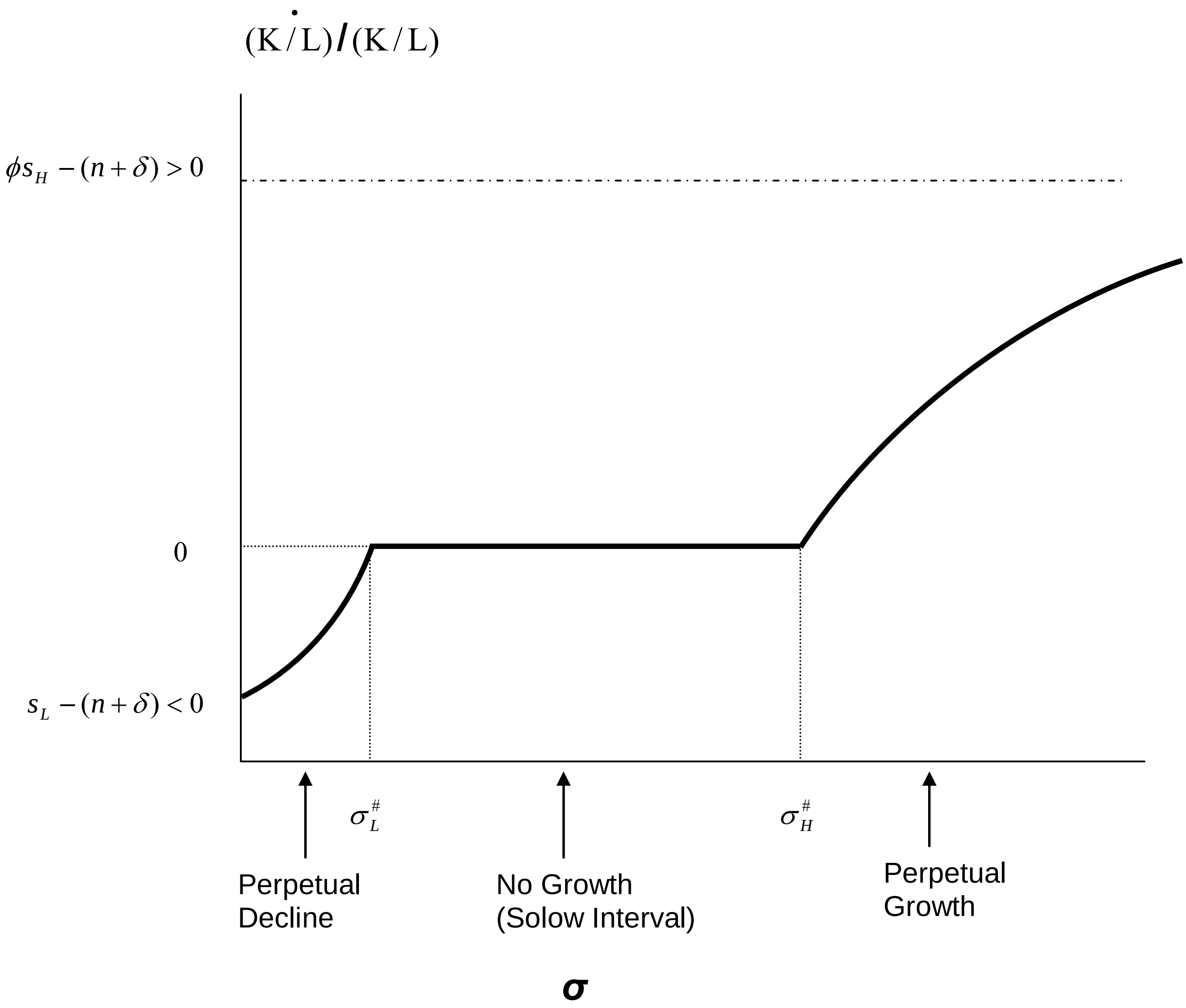
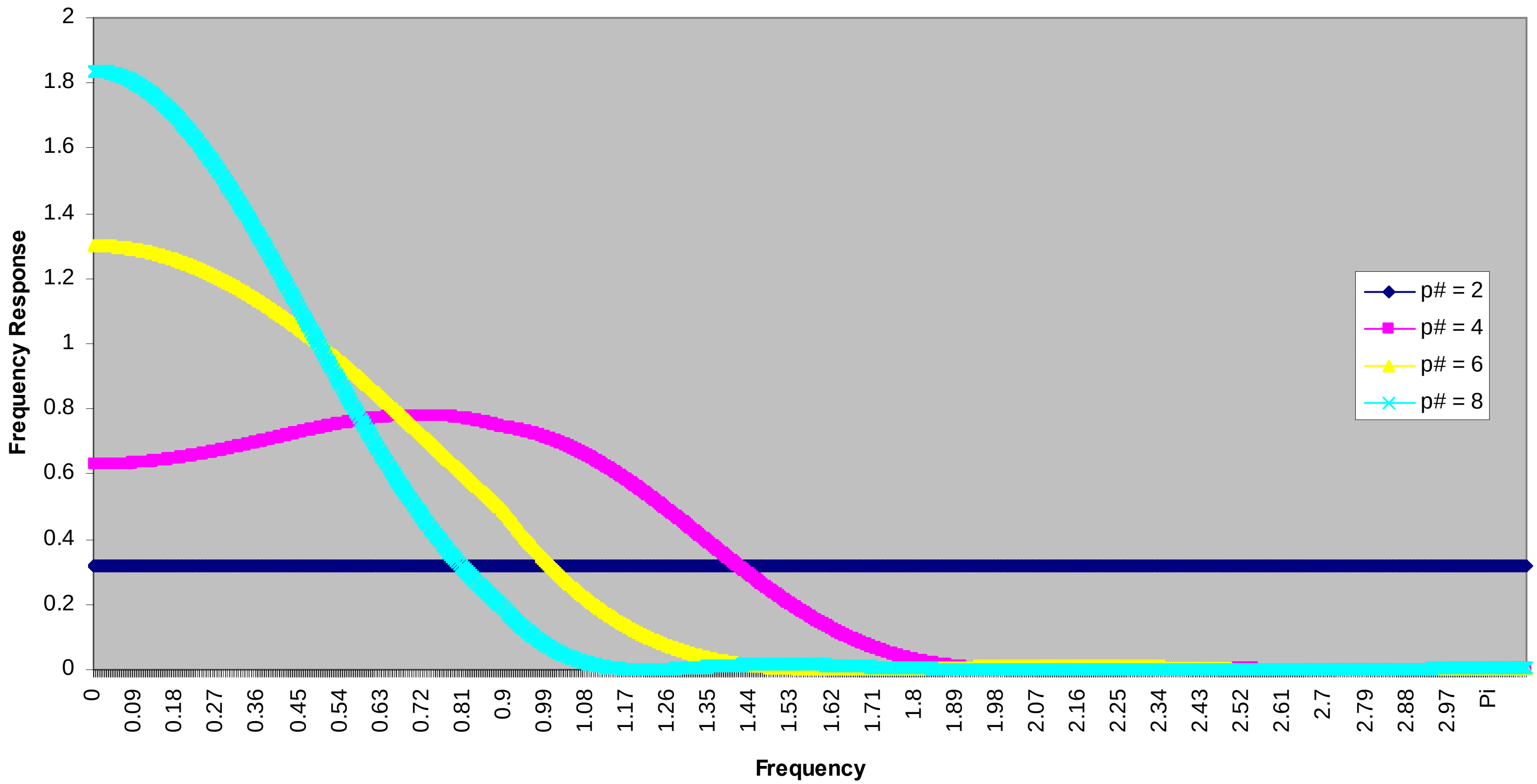


Figure 2
Frequency Response Of a[ω : p#, q]
Equation (17d)
Various Critical Periodicities (p#) With q=3



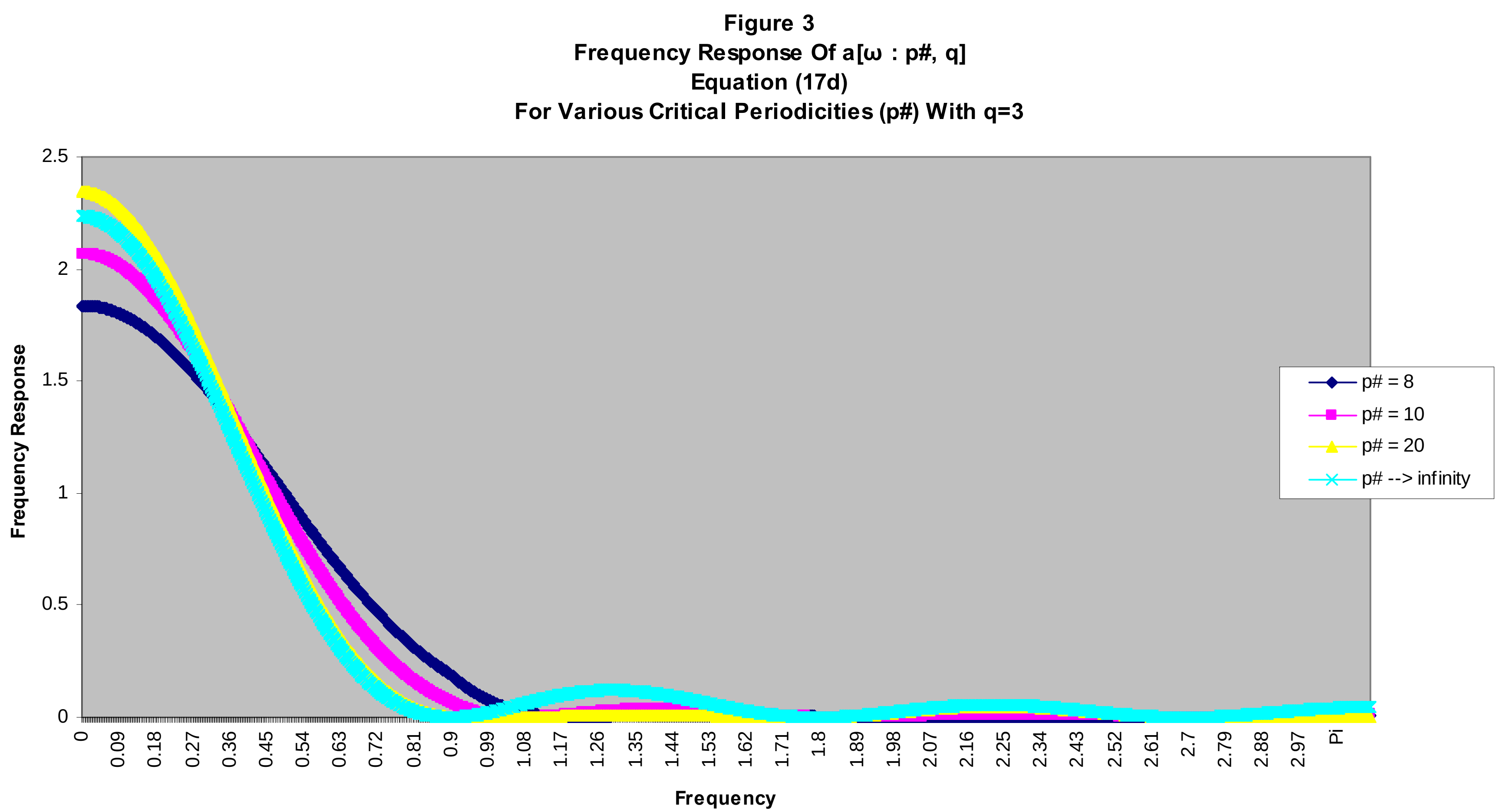


Figure 3
Frequency Response Of a[ω : p#, q]
Equation (17d)
For Various Critical Periodicities (p#) With q=3
Frequency Response
2.5
2
1.5
1
0.5
0
0
0.09
0.18
0.27
0.36
0.45
0.54
0.63
0.72
0.81
0.9
0.99
1.08
1.17
1.26
1.35
1.44
1.53
1.62
1.71
1.8
1.89
1.98
2.07
2.16
2.25
2.34
2.43
2.52
2.61
2.7
2.79
2.88
2.97
Pi
Frequency
p# = 8
p# = 10
p# = 20
p# --> infinity

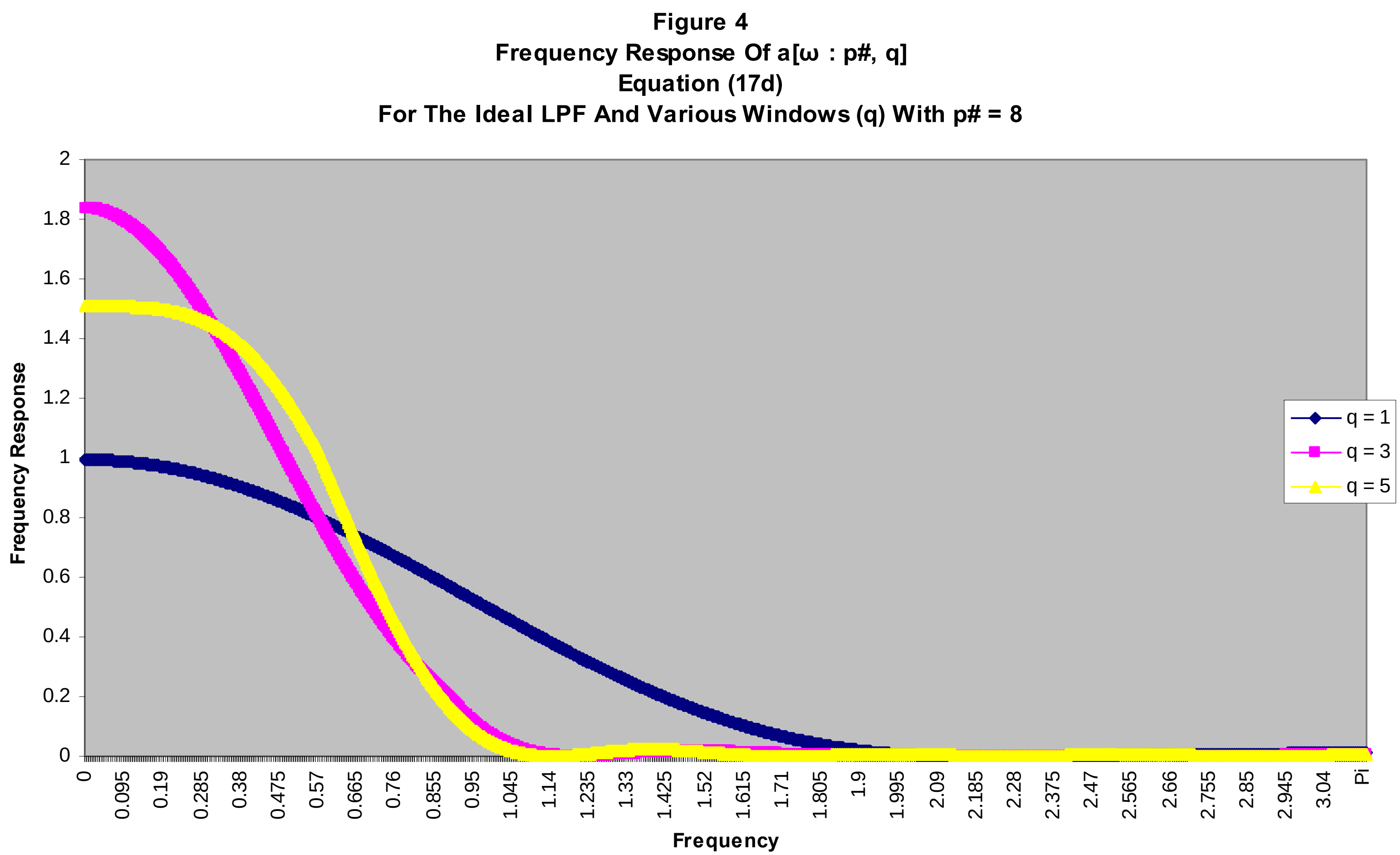


Figure 4
Frequency Response Of a[ω : p#, q]
Equation (17d)
For The Ideal LPF And Various Windows (q) With p# = 8
Frequency Response
Frequency
q = 1
q = 3
q = 5
2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
0.095
0.19
0.285
0.38
0.475
0.57
0.665
0.76
0.855
0.95
1.045
1.14
1.235
1.33
1.425
1.52
1.615
1.71
1.805
1.9
1.995
2.09
2.185
2.28
2.375
2.47
2.565
2.66
2.755
2.85
2.945
3.04
Pi