The Substitution Elasticity, Factor Shares, Long-Run Growth, And The Low-Frequency Panel Model

Robert S. Chirinko*

And

Debdulal Mallick

December 2015

The Substitution Elasticity, Factor Shares, Long-Run Growth, And The Low-Frequency Panel Model

Abstract

The value of the elasticity of substitution between labor and capital (σ) is a "crucial" assumption in understanding the secular decline in the labor share of income and long-run growth. This paper develops and implements a new strategy for estimating this crucial parameter by combining a low-pass filter with panel data to identify the low-frequency/long-run relations appropriate to production function estimation. Using spectral analysis, we assess the extent to which our choices of the critical periodicity and window defining the low-pass filter are successful in emphasizing long-run variation.

The empirical results are based on the comprehensive panel industry dataset constructed by Dale Jorgenson and his research associates. Our preferred estimate of σ is 0.40. We document that standard estimation methods, which do not filter-out transitory variation, generate downwardly biased estimates. As high frequency variation is introduced into the model variables, $\hat{\sigma}$ declines by 40% to 70% relative to the benchmark value. Despite correcting for this bias, our preferred estimate is substantially below the Cobb-Douglas assumption of $\sigma = 1$, and thus implies that the secular decline in the labor share of income cannot be explained by secular increases in the capital/income ratio or secular decreases in the relative price of investment or capital taxation.

Not for Publication

**The Substitution Elasticity, Factor Shares, Long-Run Growth, And
The Low-Frequency Panel Model**

Table Of Contents

The Substitution Elasticity, Factor Shares, Long-Run Growth, And The Low-Frequency Panel Model

> The relevant question [regarding secular movements in income shares] is whether the elasticity of substitution between labor and capital is greater or less than one.
>
> Piketty (2014a, p. 217)

> The crucial question goes to what is technically referred to as the elasticity of substitution. With 1 percent more capital and the same amount of everything else, does the return to a unit of capital relative to a unit of labor decline by more or less than 1 percent?
>
> Summers (2014)

> 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 elasticity of substitution between capital and labor, σ, is a "crucial" parameter. It affects the qualitative and quantitative answers to a host of economic questions. This paper begins by examining the role of σ in analyzing factor income shares and long-run economic growth. It then develops and implements a new strategy for estimating this crucial parameter by combining a low-pass filter with panel data techniques to identify the long-run relations appropriate to production function estimation.

The concept of σ was introduced by Hicks (1932) to analyze changes in the functional distribution of income.[1] His key 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). Recent important works by Karabarbounis and Neiman (2014) and by Piketty (2014a) and Piketty and Zucman (2014) have linked the decline in the labor share of income over the past 20 years to a secular decrease in the relative price of investment and a secular increase in the capital/income ratio, respectively. For either explanation, it is crucial (as the authors and Summers above note) that the elasticity of

---

[1] This elasticity was independently introduced by Robinson (1933, p. 256). Their formulations differed and, in the second edition of his "Theory of Wages" published in 1963, Hicks (Section III, "Notes on the Elasticity of Substitution", Sub-section 1) showed that the two formulations are equivalent. Robinson's definition has proven the more convenient and durable.

substitution is greater than one. (When depreciation is considered, the critical value of σ relevant for the study of factor incomes is much greater than one.[2]) If this elasticity is unity, relative prices and capital/income ratios have no effect on the labor share. If the elasticity is less than one, then variables highlighted in their analyses should have contributed to a counterfactual increase in the labor share. The value of σ is of first-order importance for analyzing the secular movement of the income distribution.[3]

The neoclassical revolution in growth theory introduced a variable capital/output ratio in place of the Harrod-Domar knife-edge and thus placed the burden of achieving equilibrium on the properties of the production function. When σ equals unity, the capital/labor ratio (k) converges to a positive, finite value because, as k moves towards its limiting values of 0 (or ∞), the marginal product of capital and the average product of capital both tend to ∞ (or 0). (Mathematical details are presented in Appendix I.A.) With the Inada conditions satisfied, capital accumulation, as determined by Solow's fundamental equation of motion for k, converges to zero. However, when σ departs from unity, the Inada conditions may fail, and perpetual decline or perpetual growth may occur. Intuitively, when σ is low, capital and labor are "dissimilar" productive factors (Brown, 1968, p. 50). With limited substitution possibilities, reductions in capital have little positive impact on the marginal productivity of capital. For a value of σ less than a critical value that itself is less than 1.0, k declines perpetually. Conversely, values of σ greater than a critical value that itself is greater than 1.0 lead to perpetual growth, as capital and labor are very "similar" and the marginal product of capital

---

[2] The critical σ, comparable to the one reported in all econometric studies, is substantially larger when we consider capital depreciation and net factor shares. The above discussion examines income shares in terms of the gross share of capital. Several recent commentators (listed in fn. 23, especially Rognlie (2014, who examines production functions rather than factor shares), as well as Hicks (1932)), suggest that factor shares net of depreciation are the appropriate objects of study for income distribution issues: "[a] part of the gross output must be set aside to cover depreciation in capital, and the size of this depreciation allowance, which is, of course, debited against capital's share, will affect appreciably the relative shares of the factors" (p. 297). The impact of depreciation may have been particularly important in recent years with increases in rapidly-depreciating IT capital. Defining the capital share net of depreciation, replacing the marginal product of capital with its definition in equation (II-5) in Appendix II, and differentiating with respect to the capital/output ratio, we find that the critical σ rises from 1.00 (for income shares defined gross of depreciation) to a ratio greater than 1.00 (for income shares defined net of depreciation), where this ratio is gross to net income shares. In 2013 for the corporate sector, the ratio and hence the critical σ is 1.68.

[3] The returns to factors of production in an international context also depend on σ (Jones and Ruffin, 2008; Chari, Henry, and Sasson, 2012, equation (9)).

declines very slowly.[4] Balanced growth occurs for values of σ between these two critical values. In this "well-behaved" no growth case, capital accumulation converges to zero and k to a positive, finite value. Thus, in the Solow growth model, the Harrod-Domar knife-edge is replaced by the "Solow Interval," the latter defined by the range of σ's between these two critical values. Per Solow's quotation above, the value of σ becomes a "crucial" assumption. Apart from the possibility of perpetual growth or decline, a variety of other issues in growth theory (discussed in Appendix I, Sections B to E) -- the level of steady-state income per capita, the speed of convergence, the plausibility of the Solow growth model, and the role of technical change -- depend on the precise value of σ.

Section I develops a strategy for estimating σ that applies a low-pass filter to panel data. Production function parameters are recovered by focusing on the long-run relations among 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 is not stated in terms of the raw observed values of income, but rather in terms of its permanent component. 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. Our approach also relies on permanent components, but they are extracted with a low-pass filter (a special case of the band-pass filter of Baxter and King, 1999) that affords a more general way for identifying these unobservable variables and a more transparent connection to the economic concept of the long-run.

Section II examines the theoretical properties of the spectral representation of the low-pass filter used in this study. This filter is defined in terms of two parameters -- the critical frequency ($\omega^{\#}$) defining the long-run and a window (q) for the number of lags and leads used to approximate the ideal low-pass filter. We vary these two parameters to assess the sensitivity of

---

[4] Solow (1956, pp. 77-78), Pitchford (1960), and Akerlof and Nordhaus (1967) were the first to note the possibility of perpetual growth, though they did not consider the possibility of perpetual decline. More recent analyzes of this issue can be found in Barro and Sala-i-Martin (1995, Section 1.3.3), Klump and Preissler (2000), Klump and de La Grandville (2000), de La Grandville (2009), de La Grandville and Solow (2004), and Mallick (2010).

the allocation of variance across low and high frequencies and verify that the transformed data reflect long-run variation useful in estimating production function parameters.

Sections III, IV, and V contain empirical results based on the comprehensive panel of U.S. industry dataset constructed by Dale Jorgenson and his research associates. Our econometric model relates the growth rate in the long-run capital/output ratio to the growth rate in the long-run relative price of capital (conditioned on a set of fixed time effects). The benchmark estimate of σ is 0.406 for a definition of the long-run in terms a periodicity of eight years or greater. (Note that the periodicity is inversely related to frequency, is defined as the length of time required for a series to repeat a complete cycle, and allows for a straightforward interpretation of the long-run.) We document that this result is robust to variations in the window. Moreover, there is great value in using the low-pass filter to extract permanent components. As the periodicity declines from eight to the minimum value of two years, the estimated σ declines by 40% to 70% owing to the distorting effects of transitory variation due to various frictions. At this minimum value, the low-pass filter is neutral, the raw data are not transformed, low-frequency variation is not emphasized, and estimates of σ are substantially biased downward in standard econometric models.

Section IV documents the robustness of the benchmark result to a wide variety of issues, some of which may also bias our estimate of σ: stationarity of the model variables, cross-sectional dependence in the error term, strong separability of the production function, split-samples, and measurement error in the regressor. We pay particular attention to the issue of identification in the face of technology shocks and also estimate several models that control for this important distorting factor.

Section V exploits the panel feature of our dataset and allows the estimated σ's to differ across industries. The aggregated σ corresponding to these industry estimates is larger than the benchmark estimates but well below unity. We then examine the time variation in the aggregated σ due to shifts in industry weights and relate this variation to the secular trend in the labor share. Lastly, we examine the effect on the aggregated σ of production shifts toward industries favored in a post-industrial economy.

Section VII summarizes and concludes.

I. Estimation Strategy

A. *The First-Order Condition*

Our approach focuses on long-run production relations and low-frequency variation in 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 costly frictions. This focus allows us to ignore short-run adjustment frictions that are difficult to model and may bias estimates if misspecified. Production for industry i at time t is characterized by the following Constant Elasticity of Substitution (CES) technology that depends on long-run values (*'s),

$$(1) \qquad\qquad Y_{i,t}^* = Y[K_{i,t}^*, L_{i,t}^*, X_{i,t}^*, A_{i,t}^*]$$

$$= A_{i,t}^* \left\{ \phi(K_{i,t}^*)^{[(\sigma-1)/\sigma]} + (1-\phi-\psi)(L_{i,t}^*)^{[(\sigma-1)/\sigma]} + \psi(X_{i,t}^*)^{[(\sigma-1)/\sigma]} \right\}^{[\sigma/(\sigma-1)]},$$

where $Y_{i,t}^*$ is long-run real output, $A_{i,t}^*$ is long-run technical progress, $K_{i,t}^*$ is the long-run real capital stock, $L_{i,t}^*$ is the long-run level of labor input, $X_{i,t}^*$ is the long-run level of a third factor of production (e.g., intangible capital), ϕ and ψ are distribution parameters, and σ is the elasticity of substitution between factors of production. Equation (1) is homogeneous of degree one in $K_{i,t}^*$, $L_{i,t}^*$, and $X_{i,t}^*$ and has three desirable features for the purposes of this study. First, the CES function is strongly separable and thus can include many additional factors of production (one or more X's) without affecting the estimating equation derived below. This feature is discussed further in Section IV.C and 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). Second, the Cobb-Douglas production function is a special case of the CES; as $\sigma \to 1$, equation (1) becomes $Y_{i,t}^* = A_{i,t}^* \left\{ K_{i,t}^{*[\phi]} L_{i,t}^{*[1-\phi-\psi]} X_{i,t}^{*[\psi]} \right\}$. Third, the estimating equation for capital following from this production function depends on only two parameters -- ϕ the distribution parameter and, most importantly, σ representing substitution possibilities between the factors of production.

Constrained by the CES production function (1), a profit-maximizing firm chooses capital so that its marginal product equals the relative price of capital (or 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 tax-adjusted price of output, $P_{i,t}^{Y*}$. (The firm also sets the marginal product of labor equal to the tax-adjusted nominal wage rate, $P_{i,t}^{L*}$, divided by $P_{i,t}^{Y*}$; this condition will be utilized in Section IV.F.) Differentiating equation (1) with respect to capital and rearranging terms (details are in Appendix II), we obtain the following factor demand equation for the long-run capital/output ratio,

$$(2) \qquad (K_{i,t} / Y_{i,t})^* = \phi^\sigma \, ((P_{i,t}^K / P_{i,t}^Y)^*)^{-\sigma} \, A_{i,t}^{*[\sigma-1]} \ .$$

Equation (2) shows that the impact of technical change depends on the value of σ relative to one. In our econometric model, the effects of technical change are captured in a two-way error component model of the error term,

$$(3) \qquad A_{i,t}^{*[\sigma-1]} = \exp[u_i^{KY} + u_t^{KY} + u_{i,t}^{KY}],$$

where $u_{i,t}^{KY}$ may have a non-zero mean. Inserting equation (3) into (2), taking logs of the resulting equation, and defining $ky_{i,t}^* \equiv \ln(K_{i,t} / Y_{i,t})^*$ and $p_{i,t}^{KY*} \equiv \ln(P_{i,t}^K / P_{i,t}^Y)^*$, we obtain,

$$(4) \qquad ky_{i,t}^* = \sigma \ln(\phi) - \sigma p_{i,t}^{KY*} + u_i^{KY} + u_t^{KY} + u_{i,t}^{KY} \ .$$

Removing fixed industry effects by first-differencing and defining $\tau_t^{KY} \equiv \Delta u_t^{KY}$ and $e_{i,t}^{KY} \equiv \Delta u_{i,t}^{KY}$, we obtain the following estimating equation,

$$(5) \qquad \Delta ky_{i,t}^* = \zeta^{KY} - \sigma \Delta p_{i,t}^{KY*} + \tau_t^{KY} + e_{i,t}^{KY}$$

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. As we will see in Table 1 (column 6), our dataset contains a great deal of industry-specific variation essential for estimating σ.

Given observations on $\Delta ky_{i,t}^*$, and $\Delta p_{i,t}^{KY*}$, equation (5) provides a rather straightforward framework for estimating σ. Consistent estimates are obtained because industry-specific factor prices are largely driven by aggregate factors and hence are plausibly assumed exogenous. (This assumption is relaxed in Sections IV.E, IV.F, and IV.G which employ alternative specifications

and instrumental variables to addresses the endogeneity issue.) Panel data and the two-way error component model (equation (3)) allow us to control for the effects of technical change in a general manner.[5] (Further consideration of the effects of technology shocks are presented in Sections IV.F and IV.G.) The key unresolved issue is the unobservability of long-run values denoted by *'s, an issue to which we now turn.

B. Low-Pass Filters and Long-Run Values

Previous research on capital formation has addressed this unobservability problem in several ways.[6] The cointegration approach introduced by Caballero (1994, 1999) 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, 2011). Karabarbounis, and Neiman (2014) also focus on long-run variation, and they introduce a novel method that relies on the cross-country variation in the exponential trends of income shares and the price of capital. These deterministic trends are "capturing movements from an initial to a final steady state" (p. 86). Their approach, as well as that of Caballero, defines the long-run at only the 0th frequency,[7] as opposed to a broader band of frequencies used in the macroeconomics literature (cf., fn. 9), thus discarding information. Moreover, it is desirable to decompose data frequencies with a procedure that is mutually exclusive and exhaustive. When the long-run is defined only at the 0th frequency, the spectral decomposition of variances contains an unfortunate gap between this long-run frequency ($\omega = 0.0$) and business cycle frequencies (usually in the range $\omega = \{0.79, \pi\}$). Chirinko, Fazzari, and Meyer (2011) also emphasize cross-section variation. For a panel of firms, they divide the sample period in half, average the data in each of the two intervals, difference the interval-

---

[5] The conventional procedure for modeling technical change specifies $A_{i,t} = \exp[\lambda t]$ and estimates λ as the coefficient on a linear time trend (e.g., Antras, 2004, p. 19). With this specification of $A_{i,t}$, the λ parameter would be absorbed in the constant in equation (5), and time fixed effects would capture additional variation due to technical change. Note that capital and labor augmenting technical change would enter equation (5) in a manner observationally equivalent to $A_{i,t}$.

[6] Engle and Foley (1975) also use a band-pass filter to study capital formation. They 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.

[7] For the frequency domain properties of deterministic and stochastic trends, see Granger (1964, p. 130) and Engle and Granger (1987, p. 253), respectively.

averaged data to remove nuisance parameters, and then estimate σ in a cross-section regression. A shortcoming of this estimation strategy is that it discards information in the panel dataset by relying on disjoint intervals, as well as measuring the long-run at only the 0th frequency. The most frequently used approach estimates an investment equation that begins with the first-order condition for capital and links changes in the observed capital stock to changes in the unobserved long-run capital stock by assuming that the latter is determined by changes in output and the relative price of capital and that these changes are distributed over time due to various short-run frictions (Chirinko, 1993, Section II). Relying on investment data solves the unobservability problem by imposing a set of assumptions about dynamics and frictions affecting the investment process. These assumptions – for example, convex and non-convex adjustment costs or the existence and nature of financial frictions -- can be controversial.

Our approach also focuses on the first-order condition for capital that holds in the long-run but uses a low-pass filter (LPF) to measure the long-run values of variables denoted by *'s.[8] An LPF allows frequencies lower than some critical frequency, $\omega^{\#}$, to pass through to the transformed series but excludes frequencies higher than $\omega^{\#}$. Baxter and King (1999) present two important results regarding low-pass filters 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, $x_t^*[\omega^{\#}, q]$, where q represents the "window," the length of the lags and leads used in computing the filter,

(6a)
$$x_{i,t}^*[\omega^{\#}, q] = \lim_{q \to \infty} \sum_{h=-q}^{q} d_h[\omega^{\#}] x_{i,t-h}$$

(6b)
$$d_h[\omega^{\#}] = d'_h[\omega^{\#}] + \theta[\omega^{\#}, q]$$

(6c)
$$d'_h[\omega^{\#}] = \omega^{\#}/\pi, \quad h = 0,$$

[8] Lucas (1980, especially fn. 1 and p. 1013) and Summers (1983) also use low-frequency filtered data to measure long-run variables. Relative to Lucas' exponentially smoothing filter, the LPF used in our paper can more easily be related to periodicities and readily interpretable magnitudes. Apart from Lucas' filter, there are two additional filters for isolating frequencies of interest (to the best of our knowledge, neither has been used to estimate σ). The Hodrick-Prescott (HP) filter (1997) has been primarily used to study business cycle frequencies in quarterly data. Baxter and King (1999, Section V.C) note two problems with applying the HP filter to annual data: the appropriate value for the smoothing parameter (λ) is not well established and there is a substantial divergence between HP and band-pass filters. The Christiano-Fitzgerald (2003) filter is based on the assumption that the raw data follow a random walk. This assumption is important in deriving the estimator (cf. their fn. 17) and is not appropriate for the data in the present study.

(6d) $$d'_h[\omega^\#] = \sin[|h|\,\omega^\#]/(|h|\,\pi), \quad h = \pm 1, \pm 2, \ldots, q,$$

(6e) $$\theta[\omega^\#, q] = \lim_{q \to \infty}\left(1 - \sum_{h=-q}^{q} d'_h[\omega^\#]\right)\Big/(2q+1),$$

(6f) $$\omega^\# = 2\pi/p^\# = h[p^\#], \quad p^\# = [2, \infty),$$

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 (6c) and (6d)) and a frequently imposed normalization that the $d_h[\omega^\#]$'s sum to 1 (per the constant $\theta[\omega^\#, q]$ computed in equation (6e)). Equation (6f) 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 somewhat easier to interpret than frequencies, this relation will prove useful when discussing the empirical results.

As $q \to \infty$, we obtain the ideal LPF for x_t, thus requiring an infinite amount of data and posing a challenge to implementing equation (6). Baxter and King's second important result is that the *optimal approximate* LPF for a window of finite length q truncates the symmetric moving average at q. Thus, for $|h| \le q$, the $d_h[\omega^\#]$'s are given in equations (6); for $|h| > q$, $d_h[\omega^\#] = 0$. The optimal approximate LPF for the critical frequency $\omega^\#$ (or, equivalently, the critical periodicity $p^\#$) and window q, $LPF[\omega^\#, q]$, is given by equations (6) for any finite q.

II. Spectral Properties of the Low-Pass Filter

Our estimation strategy is designed to emphasize long-run variation, and this section uses spectral analysis to evaluate the extent to which our approach is successful and its sensitivity to the choices of $\omega^\#$ and q. (Details about the spectral analysis used in this sub-section are provided in Appendix III.) The properties of a time series can be conveniently represented in the frequency domain by the spectrum defined as a Fourier transformation (a series of trigonometric functions) of autocovariances associated with a time series. The spectrum is defined over frequencies (ω) from 0 to π, and the integral of the spectrum from 0 to an arbitrary frequency

($\omega^{\#}$) less than or equal to π measures the variance of the series over those frequencies. This property will be particularly useful for assessing how well our econometric model emphasizes long-run variation.

Transformations of the data alter the spectra and hence the weights given to long-run and short-run frequencies. To compute the spectrum of a transformed series, we rely on the fundamental result from spectral analysis linking the spectrum of an output series

$\left(g_{x*}[\omega], \ x_{i,t}^{*} = \left\{ (K_{i,t} / Y_{i,t})^{*}, (P_{i,t}^{K} / P_{i,t}^{Y})^{*} \right\} \right)$ to the product of the spectrum of a raw input series

$\left(g_{x}[\omega], \ x_{i,t} = \left\{ (K_{i,t} / Y_{i,t}), (P_{i,t}^{K} / P_{i,t}^{Y}) \right\} \right)$ and a scalar $\left(\alpha[\omega, \omega^{\#}, q] \right)$ that is nonnegative, real, and,

in general, a function of ω, $\omega^{\#}$, and q. For $LPF[\omega^{\#}, q]$ transformed data defining long-run values (equations (6)), the spectrum is defined over the interval $\omega = [0, \pi]$ as follows,

(7a) $$g_{x*}[\omega, \omega^{\#}, q] = \alpha[\omega, \omega^{\#}, q] \ g_{x}[\omega],$$

where $x_{i,t}^{*}$ and the scalar associated with the LPF are defined in equation (6a) and equation (III-3a) in Appendix III, respectively. In addition to the LPF transformation, the variables entering the estimating equation derived in Section I depend on two other transformations that take logarithms of the long-run values defined by the LPF (equation (4); $y_{i,t}^{*} \equiv \ln[x_{i,t}^{*}]$) and first-difference the logarithmic variables to remove industry fixed effects (equation (5); $z_{i,t}^{*} \equiv \Delta y_{i,t}^{*}$). The spectra associated with these transformations are as follows,

(7b) $$g_{y*}[\omega, \omega^{\#}, q] = \beta \ g_{x*}[\omega, \omega^{\#}, q],$$

(7c) $$g_{z*}[\omega, \omega^{\#}, q] = \gamma[\omega] \ g_{y*}[\omega, \omega^{\#}, q].$$

where the scalars associated with the logarithmic and first-difference transformations are defined in equations (III-3b) and (III-3c) in Appendix III, respectively. Thus, the spectra for the z* variables, which are the ones that enter the estimating equation and have been subject to three transformations, are given as follows,

(7d) $$g_{z*}[\omega, \omega^{\#}, q] = \alpha[\omega, \omega^{\#}, q] \ \beta \ \gamma[\omega] \ g_{x}[\omega].$$

An important property of equation (7d) is that the spectrum of z* is independent of the ordering of the transformations because the scalars representing these transformations enter multiplicatively.

Equation (7d) allows us to examine the extent to which our estimation strategy emphasizes long-run frequencies. Since the spectra for the raw series ($g_x[\omega]$) and the scalars associated with the logarithmic and first-difference transformations (β and $\gamma[\omega]$, respectively) do not depend on $\omega^{\#}$ or q, they will not affect relative comparisons. Alternative values of $\omega^{\#}$ or q, will only affect $\alpha[\omega, \omega^{\#}, q]$, the scalar associated with the $LPF[\omega^{\#}, q]$. To ensure comparability in the analyses to follow that vary $\omega^{\#}$ and q, the areas under the spectra from 0 to π are normalized to one by an appropriate choice of normalizing constants (see equations (III-3) in Appendix III for details).

Our first set of analyses holds the window fixed at q = 3 and examines different values of the critical frequency, $\omega^{\#}$, that determines which frequencies are passed-through by the $LPF[\omega^{\#}, q]$. Recall that, per equation (6f), there is a one-to-one inverse relationship between frequencies (ω and $\omega^{\#}$) and periodicities (p and $p^{\#}$), and it is easier to discuss the properties of low-pass filters in terms of the latter. Four values of $p^{\#}$ are considered in Figure 1. We begin with the minimum value of the critical periodicity, $p^{\#} = 2$, which corresponds to either a standard investment equation that does not transform the raw data (other than the logarithmic and differencing operations) or an untransformed first-order condition. The associated frequency response is flat, indicating that this estimator does not reweight the variances of the raw series across frequencies. By contrast, our benchmark model results in a substantial reweighting. With $p^{\#} = 8$, the benchmark model emphasizes the variances from periodicities greater than or equal to eight years (corresponding to $\omega^{\#} \leq 0.79$ on the horizontal axis), thus allocating a substantial amount of weight to those frequencies appropriate for estimation of production function parameters. The remaining entries in Figure 1 are for the intermediate cases, $p^{\#} = 4$ and $p^{\#} = 6$.

The benchmark model is based on the assumption that periodicities greater than or equal to eight years contain useful information for the parameter estimates. How much additional

reweighting occurs when the critical periodicity is greater than eight years? We examine values of $p^{\#}$ equal to 10, 20 and, in the limit, ∞. The $LPF[\omega^{\#},q]$'s corresponding to these critical values are graphed in Figure 2. 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 result is consistent with what appears to be a well-accepted standard for separating long-run frequencies from short-run frequencies.[9] Second, results in Figure 2 suggest that parameter estimates are likely to be insensitive to the critical periodicity for values of $p^{\#}>8$.

We can also use equation (7d) 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 --> ∞. This ideal filter, $LPF[\omega^{\#},q\to\infty]$, is represented by the rectangle in Figure 3 that only passes-through frequencies less than 0.79; frequencies greater than 0.79 are totally excluded by this ideal filter. 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 transformed model variables. (This unwanted variation can be viewed as measurement error and will be analyzed in Section IV.E.) However, increasing q is costly in terms of lost degrees of freedom.

This tradeoff between approximation error and degrees of freedom is assessed in Figure 3, which plots $LPF[\omega^{\#},q]$ for $\omega^{\#}=0.79$ (corresponding to $p^{\#}=8$) 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. We measure this contamination by $\chi[q]$, defined as the area under the frequency response curve for the interval $\omega=[0.79,\pi]$ (the "unwanted" frequencies). For q = 1, $\chi[q=1]=0.314$. As q increases to 3 and then 5, contamination is

---

[9] A critical periodicity of eight years 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).

greatly reduced with $\chi[q=3] = 0.055$ and $\chi[q=5] = 0.054$, respectively. The $\chi[q]$'s for q = 3

and q=5 are nearly identical. Since using a window of q = 5 is costly in terms of lost degrees of

freedom and roughly the same frequencies are emphasized, we will adopt q = 3 as our preferred

window, though robustness will be examined with q = 1 and q = 5.[10]

III. Benchmark Empirical Results

This section estimates the substitution elasticity using our low-frequency panel 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

1960-2005.[11] Summary statistics are presented for the capital/output ratio in Panel A of Table 1.

The first and third rows are the model variables; the second and fourth rows are comparable

variables that have not been transformed by the LPF. Comparisons of the first row to the second

row and then of the third row to the fourth row document that the LPF substantially reduces

variation in the model variables. Columns 3 and 4 for the interquartile range and column 5 for

the within industry standard deviation show that variation is approximately halved due to the

reweighing of the lower frequencies. Column 6 reports the variance due to industry-specific

time variation as a percentage of the overall variance in the series and confirms that, even with

industry and common time variation removed (see Table Notes for details), there remains a great

deal of variation within industries with which to estimate the parameter of interest. Panel B

reports comparable statistics for the capital/labor ratio, which will be analyzed in Section IV.F.

The benchmark OLS results from our low-frequency panel model based on $p^{\#} = 8$ and

q = 3 are as follows,

(8) $$\Delta ky_{i,t}^{*} = \; 0.006 \; - \; 0.406 \; \Delta p_{i,t}^{K*} + \tau_t^K + e_{i,t}^{K*}. \qquad R^2 = 0.503$$
$$\qquad\qquad [0.003] \quad\; [0.034]$$

---

[10] Baxter and King (1999, pp. 581-582) reach a similar conclusion concerning q based on their analysis of a high-pass filter.

[11] The data are obtained at http://scholar.harvard.edu/jorgenson/data and are described in Jorgenson and Stiroh (2000, especially the appendices). These data cover the entire U.S. economy and differ from the National Income and Product Accounts by imputing a service flow to consumer durables and by making "relatively small imputations for the service flows from owner-occupied housing and institutional equipment" (p. 187). The effective time dimension equals the 46 datapoints contained in the dataset for a given industry less 2q for the construction of the LPF less one for first differencing.

The point estimate for σ is 0.406, and it is precisely estimated. Since the low-pass filter creates overlapping observations, the standard errors [in braces] are computed with the procedure of Newey and West (1994); we use (T-1) lags, which is equivalent to using a Bartlett kernel with a bandwidth of T. This correction is important and doubles the size of the standard error relative to a White (1980) procedure. The results in equation (8) are robust in several dimensions. When important differences occur, they are due to the presence of high frequency variation in the model variables, as discussed below.

Table 2 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 robust to variations in q. For example, when $p^{\#}=8$, $\hat{\sigma}$'s are 0.331, 0.406, and 0.379 for q of 1, 3, and 5, respectively. As the window is increased, more data are used in computing the filters, less data are available for estimation, and the standard errors generally increase, especially for larger values of $p^{\#}$. Nonetheless, the standard errors for the $\hat{\sigma}$'s remain less than 0.05 for all entries. (The R^{2}'s are not strictly comparable across cells because the dependent variable depends on $p^{\#}$ and q.) Consistent with the theoretical analysis in Figure 3, these results suggest that little is gained by increasing the size of the window and compromising degrees of freedom above q = 3.

Table 2 also allows us to assess the robustness to variations in $p^{\#}$ for a given q by reading down the columns. For q = 3 in column 2, as $p^{\#}$ increase from 8 to ∞, the $\hat{\sigma}$'s hardly change. This robustness is consistent with the theoretical analysis in Figure 2 suggesting that the relevant information about the long-run has been largely captured at $p^{\#}=8$. However, 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 (which do not depend on $p^{\#}$; cf. eqns. (6) and (7)). In this case, $\hat{\sigma}$ drops by over 40% relative to the benchmark value (0.229 vs. 0.406).

Our estimation strategy is based on the assumptions that short-run adjustment frictions reflected in raw data bias estimates of production function parameters, and these frictions become less important in data transformed to emphasize higher periodicities (or lower frequencies). The results discussed in the above paragraph and presented in Table 2 are consistent with these assumptions. Moreover, when we reverse our estimation procedure and use

the filter to <u>emphasize short-run variation</u> (i.e., we filter-out data with periodicities greater than $p^{\#} = 8$), the roles of adjustment frictions are magnified, and the $\hat{\sigma}$ of 0.192 [0.025] is lower than the comparable $\hat{\sigma}$'s that either emphasize long-run variation ($p^{\#} \geq 8$) or are neutral ($p^{\#} = 2$). Short-run transitory variation affects $\hat{\sigma}$ and, as has been frequently noted with permanent income models (e.g., Hayashi, 2000, pp. 194-195), transitory variation tends to attenuate point estimates. Standard estimation methods, which do not filter-out transitory variation, generate downwardly biased estimates of σ.

IV. Econometric Concerns

This section examines the sensitivity of our benchmark result in equation (8) to several issues that may bias, and in some instances bias downward, our estimates of σ: stationarity of the model variables, cross-sectional dependence, the strong separability of the production function, split samples, and measurement error. We also consider challenges to identification due to technology shocks. Tables 3 and 4 are presented in abbreviated form. The complete tables can be found in Appendix VI.

A. Stationiarity

The stationarity of the model variables needs to be considered because spectral filters are not strictly valid with nonstationary data. The Solow growth model implies that both the capital/output and relative price of capital series are stationary, an implication consistent with the two of the stylized facts of growth advanced independently by Kaldor (1961) and by Klein and Kosobud (1961). However, these theoretical facts and historical regularities are at odds with recent data on falling labor income shares mentioned in the introduction. Hence, the low-pass filter used in this study may not be strictly applicable.

Given this tension between theoretical predictions and recent empirical facts, it is useful to examine formally the stationarity properties of our panel data. We use a test for stationarity proposed by Pesaran (2007) that extends the standard augmented Dickey-Fuller test to allow for cross-sectional dependence in panel data. See Appendix IV for details of the testing procedure. As shown in Table IV.A in Appendix IV, we reject the null of stationarity at the 1% or 10% level for the relative price variable (depending on the inclusion or exclusion of a deterministic trend). For the capital/output ratio, the null hypothesis of non-stationarity is sustained.

In light of this evidence, we exploit the commutative property of our filtering procedure. As shown in equation (7d), the three transformations are multiplicative. Thus, we can alter the sequence to ensure that we are filtering stationary data. Given the concerns about stationarity, we first take logs and then difference. The resulting transformed data are stationary (see Table IV.B in Appendix IV). We then apply our low-pass filter to these transformed data.

The multiplicative property in equation (7d) holds only for infinitesimally small transformations.[12] We confirm that the discrepancy introduced by finite transformations is empirically unimportant. The results for a range of $p^{\#}$'s ranging from 2 to ∞ are presented in Table 3, column 1. The results prove very robust. For example, for our benchmark model with the transformation sequence LPF/logs/first-difference, $\hat{\sigma} = 0.406\,[0.034]$. When we change the order of the sequence to logs/first-difference/LPF, $\hat{\sigma} = 0.395\ [0.032]$. Thus, our benchmark results are not sensitive to the stationarity properties of our data.

B. Cross-Sectional Dependence

Our OLS results might be inconsistent in the face of cross-sectional dependence in the error term that is not fully captured by time fixed effects and hence becomes part of the error term. Correlation between these shocks and the regressors would lead to inconsistent estimates of σ. Even absent such correlation, the shocks will lead to biased standard errors. Cross-sectional dependence could be accounted for by Seemingly Unrelated Regression, but this estimator is not feasible when the number of cross-section units exceeds 10. The common correlated effects (CCE) estimator introduced by Pesaran (2006) is feasible for panels with a large number of cross-section units, and it accounts for the effects of cross-sectional dependence by including cross-section averages (CSA's) of the dependent and independent variables as additional right-hand side variables. Details of this estimation procedure are provided in Appendix V.

The results for the CCE model are presented in Table 3, column 2. The CCE model requires 35 additional estimated parameters, and the standard errors increase markedly.[13]

---

[12] In addition, "matching conditions" (essentially that the scalars are finite for all ω from 0 to π) need to be satisfied (Koopmans, 1974, p. 97).

[13] For example, for our preferred model with $p^{\#} = 8$ and q=3, the comparable White corrected standard errors are larger by a factor of 3.5. The standard errors reported in column 2 of Table 3 are not Newey-West corrected because of computational issues with STATA related to non-linearity in the parameters.

This model nests the standard time fixed effects model. While we reject the restriction in the standard time fixed effects model, for most values of the critical periodicity (except for $p^{\#} = 2$), the $\hat{\sigma}$'s are very similar to those from the benchmark model.

C. Strong Separability of the Production Function

All of the above estimates have been based on output defined by gross-value, a concept consistent with the CES production function specified in equation (1). The CES function can easily accomodate a value-added definition of output that depends only on capital and labor by setting the distribution parameter $\psi = 0$, and the associated parameter estimates can serve as a useful robustness check. Each approach is based on different, non-nested assumptions. If we use a value-added definition of output in our empirical model, we are maintaining the validity of the double deflation method whereby both gross output and intermediate inputs are deflated separately. If we use a gross-value definition of output in our empirical model, we are maintaining the validity of the strong separability assumption inherent with the CES model.[14] Estimates based on the value-added definition of output are presented in Table 3, column 3. They are higher than those from the benchmark model; for example, for our benchmark model,

$$\hat{\sigma}^{\text{Value-Added}} = 0.543[0.038] > \hat{\sigma}^{\text{Gross-Value}} = 0.406[0.034].$$

D. Split-Samples

To further assess the robustness of our results, columns 4 and 5 of Table 3 contains OLS estimates from the first and second halves of the sample, 1960-1982 and 1983-2005, respectively. The results closely follow those reported previously. For example, for our preferred specification with $p^{\#} = 8$ and q = 3, the $\hat{\sigma}$'s from the first and second halves of the sample are 0.336 and 0.460, respectively. These estimates are not economically different from our preferred estimate from the full sample of 0.406, though, owing to the precision of our estimates, these deviations from the benchmark are statistically significant.

---

[14] For a strongly separable production function, the marginal rate of transformation between factors K and L do not depend on any other factors of production. There is a long-standing debate about the production relations among capital, labor, and energy beginning with an exchange between Berndt and Wood (1981) and Griffin (1981). We prefer not to get involved in that debate and the estimation of nested CES and translog models, and instead we estimate the value-added specification as a robustness check on the parameter of primary interest in this study.

E. Measurement Error

The OLS estimates reported above are valid under the assumption that the growth rate in the long-run relative price of capital is uncorrelated with the error term. This assumption may be compromised for a variety of reasons. A common factor may exist that affects both the dependent and independent variables but only enters the specification through the error term. Technical change may be one such common factor correlated with the regressor via the relative price of investment (we will consider the role of technology shocks explicitly in Sections IV.F and IV.G).[15] Perhaps the most important source of error arises from our procedure for measuring the true long-run value. This issue is particularly relevant because, if quantitatively important, it will bias $\hat{\sigma}$ downward and away from 1.0. As discussed in the introduction, whether σ is above or below one has major implications for a variety of economic issues.

As discussed in Section II and documented in Figure 3, the LPF is subject to error relative to the ideal filter. The $\chi[q]$ statistic quantifies the extent to which our empirical measures of long-run values are "contaminated" by unwanted frequencies. The $\chi[q=3]$ statistic is only 0.055, suggesting that we are capturing most of the relevant variation.

Notwithstanding this favorable statistic, measurement error could affect the estimates, and instrumental variable (IV) estimates provide a useful robustness check. A lagged regressor is a familiar candidate instrumental variable. However, in our model, the low-pass filter depends on leads that may be correlated with the contemporaneous error term.   Hence, a lagged regressor is not a suitable instrument unless it is lagged many periods, potentially compromising its relevance. While the model developed in Section I places restrictions on the construction of model variables, it does not place restrictions on the instrument, and thus we create an alternative instrument, $p_{i,t-2}^{KY\&}$, as a one-sided filter of the relative price of capital and lag it two periods (see the note to Table 4 for details). This variable will be uncorrelated with the contemporaneous error term and will deliver consistent estimates.

---

[15] The extent to which capital augmenting technical change (CATC) enters the error term is uncertain. There is one tradition – beginning with Greenwood, Hercowitz, and Krusell (1997) – that assumes impounding into the price of investment goods is complete, and hence CATC does not enter the error term. (The Jorgenson data used in this study are constructed under the assumption of complete impounding; Jorgenson and Stiroh, 2000, p. 191.) The assumption of complete impounding is controversial (e.g., Cummins and Violante, 2002). Partial impounding would lead to the endogeneity problem addressed by the IV estimates presented in this sub-section and Sections IV.F and IV.G.

Table 4, column 1, contains IV results for equation (5). The R^2 statistic is replaced by an F-statistic for the first-stage auxiliary regression of the regressor on the instrument and time dummies. Instrumental relevance is assessed with the latter statistic proposed by Stock, Wright, and Yogo (2002); for the auxiliary regression, we compare the F-statistic for the goodness of fit to a critical value of 8.96 (reported in their Table 1). Based on this critical value, the instrument is not weak for all models except for the one in row 1, which does not transform the raw data to emphasize lower frequencies. For the other estimates with relevant instruments, the estimated σ's tend to be about 10% larger than the comparable OLS estimates. For example, for our preferred estimate based on $p^\# = 8$ and $q = 3$, $\hat{\sigma}^{IV} = 0.438 \, [0.062]$ is somewhat larger than $\hat{\sigma}^{OLS} = 0.406 \, [0.034]$. The IV standard errors are twice as large as their OLS counterparts. Nonetheless, the results continue to decidedly reject the Cobb-Douglas hypothesis that $\sigma = 1$.

The role of measurement error can also be assessed by answering the following question: how important would measurement error need to be in order to reconcile our estimates with a true value of $\sigma = 1$? This question can be answered with the following formula for the asymptotic bias of $\hat{\sigma}$ in the presence of classical measurement error,

$$(9) \qquad \left(\sigma^@ - \hat{\sigma} \right) = \left(VAR\left[\xi_{i,t} \right] / VAR\left[\Delta p_{i,t}^{KY} \right] \right) \sigma^@ ,$$

where $\sigma^@$ and $\hat{\sigma}$ are the true and estimated values of σ, respectively, $\xi_{i,t}$ is the measurement error in the regressor, and VAR[.] is the variance operator. If $\sigma^@ = 1.00$, then the measurement error variance would have to account for at least 60 percent of the variance in the benchmark model's regressor. This seems like an unreasonably large figure.

The evidence presented in this sub-section suggests that measurement error has little effect on biasing downward $\hat{\sigma}$. The result is perhaps not surprising since measurement error is usually modeled as occurring at higher frequencies, the ones that have been filtered-out by our estimation procedure.

F. Identification And Technology Shocks – Part 1

Technology shocks pose a major challenge to proper identification of σ in our benchmark specification. In the above regressions, technology shocks may well be correlated with both the regressor and the instrument. In these cases, $\hat{\sigma}$ is biased and inconsistent. We

address this important concern in this sub-section from three different angles that either eliminate technology shocks or create instruments likely to be valid with respect to technology shocks. Sub-section G represents technology shocks parametrically and assesses their effects on $\hat{\sigma}$.

First, we derive a specification that eliminates technology shocks from the estimating equation. Apart from the first-order condition for capital that is the basis for the benchmark model, the optimization problem also generates a similar first-order condition and the associated estimating equation for labor (and for other factors of production). By subtracting this equation from that for capital (equation (5)), technology shocks are eliminated, and we obtain the following estimating equation for the capital/labor ratio, $k\ell_{i,t}^{*}$,

(10) $$\Delta k\ell_{i,t}^{*} = \zeta^{KL} - \sigma \Delta p_{i,t}^{KL*} + \tau_{t}^{KL*} + e_{i,t}^{KL*},$$

where $p_{i,t}^{KL*} \equiv \ln(P_{i,t}^{K} / P_{i,t}^{L})^{*}$, and the other elements in (10) parallel those defined in equation (5). The important difference is that the technology shock, $A_{i,t}^{*}$, no longer enters the error term.

Estimates of this technology-shock-free equation are presented in Table 3, column 6, and the $\hat{\sigma}$'s are uniformly lower than the benchmark OLS estimates across all periodicities. For our preferred model with $p^{\#} = 8$ and $q = 3$, $\hat{\sigma}^{KL} = 0.333\,[0.053]$ is smaller than $\hat{\sigma}^{Benchmark} = 0.406\,[0.034]$, though the difference is not statistically significant owing to the larger standard error for the KL specification. This result is consistent with the following scenario for the benchmark model: a positive technology shock increases output supply (cf. equation (1)), decreases output price, and increases the relative price of capital (PK/PY). Since this latter variable enters the regression with a negative sign, the technology shock lowers the independent variable. The technology shock also lowers the dependent variable, the K/Y ratio (cf. equation (2) for $\sigma < 1$), and hence biases $\hat{\sigma}$ upward. Under this scenario, the KL specification, which eliminates the effects of technology shocks, should generate a lower $\hat{\sigma}$.

Second, the above scenario suggests an alternative instrument that will be less affected by technology shocks than the one used in Table 4, column 1, $p_{i,t-2}^{KY\&}$. This new instrument is the ratio of the price of capital (with suitable tax adjustments) to the price of output. As sketched in the scenario, the first-order impact of technology shocks is likely to be on the price of output. To

avoid the potentially distorting effects of technology shocks, we define a second instrument

solely in terms of the price of capital, $p_{i,t-2}^{K\&}$, that is one-sided and lagged two periods.

Estimates with this instrument are presented in Table 4, column 2. We again find that controlling for technology shocks leads to a lower $\hat{\sigma}$. For our preferred model with $p^{\#} = 8$ and $q = 3$, the estimated $\hat{\sigma}$ with this instrument is 0.345 [0.087], very close to the estimate of 0.333 [0.053] from the KL specification.

Third, all of the previous IV estimates are based on instruments that are "internal" to the firm and the objects in its optimization problem. Hence these instruments and their validity may be compromised by technology shocks in a variety of unknown ways. An alternative approach that avoids the potential problem of invalid "internal" instruments relies on "external" instruments. Federal tax policy is surely one such valid instrument.[16] However, two concerns arise with using federal tax instruments: they restrict the use of time fixed effects and they may not be relevant in the sense of being weakly correlated with the regressor. We examine the first of these by rerunning the benchmark model without time fixed effects. The OLS results presented in Table 4, column 3, confirm that time fixed effects have little impact on $\hat{\sigma}$. This robustness is not surprising since our estimation strategy emphasizes long-run variation (with the LPF), and hence the short-run variation captured by time fixed effects has little influence on $\hat{\sigma}$.

IV estimates are presented in Table 4, columns 4 to 6, with a variety of marginal effective federal tax rates on capital income: corporate firm-level, corporate total, and non-corporate total. The second concern about using these "external" instruments proves unwarranted. The F-statistics computed from the first-stage auxiliary regression of $\Delta p_{i,t}^{KY*}$ on the tax instrument(s) are again used to assess instrument relevance. The null hypothesis of a weak instrument is evaluated at the 5% level and rejected for all models using tax instruments. The results in column 4 are based on the corporate firm-level instrument (which we believe is the most accurately measured tax rate), and $\hat{\sigma}$ rises to 0.452 [0.103] relative to the benchmark value of 0.406 [0.017]. Similar results are obtained when we use all three tax rates as instruments (column 5) or these three tax rates and $p_{i,t-2}^{KY\&}$ as instruments (column 6). In these latter two

---

[16] We thank one of the referees for suggesting this tax instrument, and Jane Gravelle for providing the tax data. See the note to Table 4 for details about the three tax series.

cases, the models are overidentified, and these restrictions can be tested by a Hansen J-statistic. For models that pass through high frequencies/low periodicities (i.e., $p^{\#} = 2$ and $p^{\#} = 4$), the overidentifying restrictions are rejected at the 1% and 10% levels, respectively. These restrictions are also (nearly) rejected at the 10% level for the models defining the long-run at the 0th frequency (i.e., $p^{\#} \to \infty$). The J-tests can be interpreted as portmanteau tests of model misspecification, and they suggest the value of filtering the data over a range of long-run frequencies to generate a correctly specified model.

G. Identification And Technology Shocks – Part 2

This sub-section represents technology shocks parametrically in two distinct ways and assesses the effects on $\hat{\sigma}$. First, technology shocks have frequently been modeled in terms of time trends, and we add the following terms to the benchmark model (after removing time fixed effects): (a) linear time trend, (b) linear + quadratic time trends, and (c) a Box-Cox transformation (BCT_i). The latter has been suggested by Klump, McAdam, and Willman (2007, p. 185) and involves adding the following term to the benchmark equation,

$$(11) \qquad\qquad\qquad BCT \equiv \lambda \left(t^{\zeta} - 1 \right) / \zeta,$$

where λ is parameter to be estimated and ζ a curvature parameter dictating the shape of technical progress. To avoid the complications associated with estimating a model non-linear in the parameters and given that our primary focus is estimating σ, we pre-set ζ to the 0.1, 0.5, and 1.5 (note that a value of 1.0 corresponds to the linear trend estimated in (a)). We repeat these five estimates with industry-specific linear and linear + quadratic time trends and replace λ by λ_i in the three BCT models. In all ten cases, the effects on $\hat{\sigma}$ are very small, with the change being less than 0.03 (in absolute value) from the benchmark estimates.

Second, an additional direct approach for measuring technology shocks relies on the parsimonious parameterization of the CES production function that is the basis of our estimates. The value-added form of this function depends on two unknown parameters and the unknown technology shock ($A_{i,t}^{*}$). Equation (1) with $\psi = 0$ can be written as the following value-added production function ($Y^{VA}[.]$) that implies an equation for the technology shock,

(12)
$$Y_{i,t}^* = Y^{VA}\left[K_{i,t}^*, L_{i,t}^*, A_{i,t}^* : \sigma', \phi[\sigma']\right] \rightarrow \tilde{A}_{i,t}^*[\sigma'] = F[\sigma']$$

where $F[.]$ is function that incorporates all known variables, and σ' represents two pre-set

values of $\sigma' = \{0.5, 1.5\}$ and differs from the estimated σ, $\hat{\sigma}[\sigma']$, which is now a function of σ'.

(Note that the technology shock does not enter the estimating equation when $\sigma' = 1.0$; cf.

equation (3).) We determine ϕ as a function of the capital income share,[17] solve for $\tilde{A}_{i,t}^*[\sigma']$,[18]

and then include the imputed technology shock as an additional regressor (entered in log

differences) in the benchmark specification. When σ' is preset to 0.50, the $\hat{\sigma}$ declines from

$\hat{\sigma}^{benchmark} = 0.406$ to $\hat{\sigma}^{\sigma'=0.50} = 0.330$. Given the ad hoc nature of the computation of

$\tilde{A}_{i,t}^*[\sigma']$, we do not want to overemphasize the point estimates. Nonetheless, this result is

consistent with those presented previously -- accounting for the effects of technology shocks

tends to lower $\hat{\sigma}$. When the technology shock is constructed under the assumption that

$\sigma' = 1.50$, $\hat{\sigma}^{\sigma'=1.50} = 0.731$ is still well below unity.

On balance, the evidence presented in Sections IV.F and IV.G suggests that technology

shocks do not have a major impact on our estimates of σ. This somewhat surprising result may

be due to the specification of the benchmark model where the effects of industry-specific and

aggregate components of technology shocks are accounted for by first-differencing and time

fixed effects, respectively.

V. Heterogeneous Industry σ_i's

Given our interest in the impact of aggregate σ on factor income shares and long-run

economic growth, the homogeneity assumption imposed across industries in the prior two

sections is a natural way of obtaining an aggregate substitution elasticity. From an estimation

perspective, however, it might be desirable to exploit the panel feature of our dataset and allow

---

[17] This result is consistent with the normalization proposed by Herrendorf, Herrington, and Valentinyl (2015; see their equation (21)). A similar result holds in the normalization proposed by Klump, McAdam, and Willman (2012, Section 2.4) at the point of normalization. The first-order equation for capital associated with either normalized CES production function is observationally equivalent to the one used in this paper (i.e., equation (4)).

[18] There is a great deal of industry-specific variation in the $\tilde{A}_{i,t}^*[\sigma']$'s; when industry and time fixed effects are removed, the remaining variation (as a percentage of overall variation) is 0.30 and 0.48 for $\sigma' = \{0.5, 1.5\}$, respectively.

the σ's to differ across industries. The statistics for industry-specific time variation (Table 1, column 6) indicate that there is sufficient variation to identify industry-specific σ's. The aggregated σ (σ_{agg}) is a weighted average of the industry σ's (σ_i's, see Appendix VII for derivation),

$$(13) \qquad \sigma_{agg} \equiv \sum_i \sigma_i * w_i, \qquad w_i \equiv \left(\sum_t K_{i,t} \Big/ \sum_t \sum_i K_{i,t} \right) = \left(K_i / K_{agg} \right)$$

where K_{agg} is the aggregate capital stock and the w_i's are industry weights defined in terms of industry capital ratios. Alternative aggregation procedures are discussed in Appendix VII.

Table 5 contains estimates of the heterogeneous industry σ_i's and the corresponding $\hat{\sigma}_{agg}$'s. Panel A, column 1, shows that the $\hat{\sigma}_i$'s are precisely estimated and are all statistically far from zero and far from unity (except Finance, Insurance & Real Estate, #33). Panel B presents several $\hat{\sigma}_{agg}$'s based on different weighting schemes.[19] Column 1 uses capital weights for all 35 industries, and the $\hat{\sigma}_{agg} = 0.657$ is larger than the benchmark value of 0.406 based on a homogeneity assumption. This difference is due to the distribution of capital weights. Column 2 weights the 35 industries equally, and the $\hat{\sigma}_{agg} = 0.417$ is very close to the benchmark value of 0.406. The estimates presented in Table 5 provide additional support for the conclusion that the σ is well-below the Cobb-Douglas value of one.

These results weight the industry σ_i's by industry capital ratios computed as averages over the full sample. How much does σ_{agg} vary due to changes in industry weights over time?[20]

$$(14) \qquad \breve{\sigma}_{agg,t} \equiv \sum_i \sigma_i * w_{i,t} \ .$$

As shown in Figure 4, the time plot of $\hat{\breve{\sigma}}_{agg,t}$ is quite provocative. There is a downward trend in the first 20 years, and then a more or less monotonic increase beginning in 1980. This increase coincides with the well-documented shift in the capital share of aggregate income. The capital

---

[19] It should be noted that the $\hat{\sigma}$ from the benchmark model is also effectively a weighted-average estimate. The heterogeneous model analyzed in this appendix weights the $\hat{\sigma}_i$'s by industry capital shares, while the homogeneous model effectively weights the $\hat{\sigma}_i$'s by relative industry variances.

[20] We thank one of the referees for posing this question.

share increases in σ_{agg} (cf. Appendix VIII). What quantitative role is played by the increase in σ_{agg} in explaining the rise in the capital share of income? Unfortunately, the nearly four point increase in $\hat{\tilde{\sigma}}_{agg,t}$ has only a very modest effect on the capital income share.

The above results are based on historical production patterns. Is σ_{agg} likely to rise as production shifts toward industries favored in a post-industrial economy? Piketty (2014b, p. 39) suggests an affirmative answer: "…the elasticity of substitution between capital and labor in the future may be different (and possibly higher) than the past one." Column 3 of Panel B allows for an informal examination of this question. We identify post-industrial industries as those that are not agriculture, manufacturing, mining, and utilities. Thus, in our dataset, we focus on the following six industries: Construction (6), Transportation (28), Communications (29), Trade (32), Finance, Insurance and Real Estate (33) and Services (34). The $\hat{\sigma}_{agg}$ for these selective six industries is 0.857. These estimates support the notion that σ_{agg} is likely to rise in future years as production and hence capital weights rise in these post-industrial industries.

We close our presentation of empirical results with a reexamination of the importance of filtering and emphasizing long-run frequencies. The results in Table 5, Panel B, are based on our preferred definition of the long-run with a periodicity of 8 years or longer. Panel A, column 2, and Panel C of Table 5 report results for the same exercise with heterogeneous $\hat{\sigma}_i$'s but where the data are not filtered (i.e., p# = 2). The results change dramatically. Thirty-one of the $\hat{\sigma}_i$'s fall, and only four rise (Food and Kindred Products, #7; Tobacco, #8; Paper and Allied Products, #13; Chemicals, #15); the average increase is 0.07. When $p^{\#} = 2$, the largest $\hat{\sigma}_i$ is 0.517 for the Oil and Gas Extraction industry. As shown in Panel C, the $\hat{\sigma}_{agg}$ of 0.196 is 70% lower than the comparable estimate using a LPF with $p^{\#} = 8$. The remaining $\hat{\sigma}_{agg}$'s in Panel C are also reduced substantially relative to their counterparts in Panel B. These disaggregate results confirm the importance of filtering and the downward bias due to unwanted transitory variation.

VI. Summary and Conclusions

The elasticity of substitution between labor and capital (σ) is a "crucial" parameter. The qualitative and quantitative answers to a host of economic issues depend on the precise value of

σ. This crucial production function parameter is estimated by combining a low-pass filter with panel data to identify the long-run relations appropriate to production function estimation. Our preferred point estimate is 0.406, and it proves robust to variations in several directions, as the alternative estimates range from 0.333 to 0.543 (or 0.657 for the heterogeneous industry model).

Although we believe the empirical evidence presented here points towards a value of σ in the neighborhood of 0.40, we recognize that the matter is far from settled. Based on cross-country data, Karabarbounis and Neiman (2014, p. 64; this paper was discussed in Section I) and Piketty (2014a, p. 221) estimate σ's of 1.25 and between 1.30 and 1.60, respectively. However, also using cross-country data, Mallick (2012) obtains lower $\hat{\sigma}$'s that range widely and average about 0.30 for both six high-income and all 90 countries.

Oberfield and Raval (2014) use U.S. manufacturing plant level data to estimate σ with the first-order condition for labor in two cross-sections. They find that $\hat{\sigma}_{plants}$'s average about 0.50 (Table I). This estimate is conceptually closest to the ones presented in our paper. Moreover, the use of cross-sections emphasizes long-run variation, and thus resonates with the approach taken in our paper. The authors then aggregate the $\hat{\sigma}_{plants}$'s accounting for reallocations of economic activity among plants and across manufacturing industries. Their $\hat{\sigma}_{mfg}$ equals 0.70; our comparable estimate for manufacturing industries is 0.39 (cf. fn. 22).[21]

Two other studies using U.S. data report $\hat{\sigma}$'s that are approximately double our preferred value of 0.40. Herrendorf, Herrington, and Valentinyl (2015, Table 1) estimate a model with a production function and capital and labor first-order conditions with sectoral U.S. data. The $\hat{\sigma}$'s for the agriculture, manufacturing and service sectors are 1.58, 0.80, and 0.75, respectively.[22] Their $\hat{\sigma}_{agg}$ equals 0.84. León-Ledesma, McAdam, and Willman (2015, Table 5) employ a similar three equation system on aggregate U.S. data, and report $\hat{\sigma}$'s ranging from 0.44 to 0.72. Common characteristics in these latter two studies are the estimation of CES production functions with untransformed data. As has been shown here, the use of untransformed data can

---

[21] See Appendix VII for a discussion of the differences in the aggregation procedures used by Oberfield and Raval and in this paper.

[22] It is interesting to note that the ranking of their sectoral σ's by HHH ($\sigma_{ag} > \sigma_{mfg} > \sigma_{ser}$) differs from those in our Table 5 ($\sigma_{ag} = 0.289 < \sigma_{mfg} = 0.393 < \sigma_{ser} = 0.514$). We have no ready explanation for the different rankings, other than noting that omitted factors – land, human capital, and intangible capital – may play a role.

be problematic; determining the direction of the bias in a non-linear CES equation embedded in an equation system is difficult. Moreover, omitted factors of production – such as land, human capital, or intangible capital – that enter into the production function (the X variable in equation (1)) may create challenges to proper specification of the estimating equations. Clearly, future research needs to analyze these different estimation procedures with a common dataset or Monte Carlo model to understand the sources of these diverse estimates of σ.

In our study, we focused on two sets of economic issues for which σ is a crucial parameter. As noted in the quotations by Piketty and Summers at the beginning of this paper, whether σ is greater than or less than one is of crucial importance in the study of factor incomes. Secular increases in the capital/income ratio or secular decreases in the relative price of investment can explain the recent decline only if $\sigma > 1$. For the United States, we find no support for this assumption. This finding leaves open the question as to what factors are responsible for the much-discussed secular decline in the labor income share observed in several advanced economies (Karabarbounis, and Neiman, 2014, Figure II, p. 71; Piketty, 2014a, Figure 6.5, p. 222). One intriguing hypothesis is that a general movement toward more business-friendly policies has relaxed regulations that lower operating costs and raise markups. This movement, interacting with a decline in unionization, is a candidate explanation being explored in ongoing work.[23]

In the Solow growth model, certain values of σ can lead to unbalanced positive or negative growth. When $\sigma < 1$, it cannot be an engine of perpetual growth. However, this value of the elasticity may be an agent of perpetual decline if it is less than a critical value that depends on the rates of saving, depreciation, population growth and the share of income accruing to labor. As shown in Table 1-A in Appendix A, $\sigma = 0.406$ is much larger than the critical value for perpetual decline in the United States. Thus, our estimate of σ suggests that this elasticity is not so small as to create unbalanced negative growth for the U.S. economy. As mentioned in the Introduction and discussed in some detail in Appendix I, the resolution of several other issues concerning long-run growth depend on the precise value of σ.

Our benchmark value of σ of 0.40 suggests that the convenient and all too often

---

[23] While the value of σ is a crucial in evaluating the drivers of the declining share of labor income, other empirical issues that have arisen include the roles of housing capital and capital depreciation; see, among others, Bridgman (2014), Rognlie (2014), Summers (2014), and Auerbach and Hassett (2015) for further discussion.

conventional assumption of a Cobb-Douglas production function used in many areas of economic analysis needs to be abandoned. DSGE models that maintain that $\sigma = 1$ amplify the true effect of price movements -- including the traditional channel of monetary policy -- relative to a model based on a lower value of this elasticity. Tax simulation models using Cobb-Douglas production functions impart a similar upward bias to the effects of tax cuts. 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 directed technical change that affects factor shares and balances growth.

REFERENCES

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.

Arrow, Kenneth J., Chenery, Hollis B., Minhas, Bagicha S., and Solow, Robert M., "Capital-Labor Substitution and Economic Efficiency," *The Review Of Economics And Statistics* 43 (1961), 225-250; reprinted in *Production and Capital: Collected Papers of Kenneth J. Arrow* Vol. 5 (Cambridge: Harvard University Press, 1985), 50-103.

Auerbach, Alan J., and Hassett, Kevin, "Capital Taxation in the Twenty-First Century," *American Economic Review* 105 (May 2015), 38–42.

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

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., and Wood, David O., "Engineering and Econometric Interpretations of Energy-Capital Complementarity: Reply and Further Results," *American Economic Review* 71 (December 1981), 1105-1110.

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.

Bridgman, Benjamin, "Is Labor's Loss Capital's Gain?: Gross versus Net Labor Shares," Bureau of Economic Analysis (June 2014).

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

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.

Chari, Anusha, Henry, Peter Blair, Sasson, Diego, "Capital Market Integration and Wages," *American Economic Journal – Macroeconomics* 4 (April 2012), 102-132.

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., Fazzari, Steven M., and Meyer, Andrew P., "A New Approach to Estimating Production Function Parameters: The Elusive Capital-Labor Substitution Elasticity," *Journal of Business & Economic Statistics* 29 (October 2011), 587-594.

Chirinko, Robert S., and Mallick, Debdulal, "Cointegration, Factor Shares, and Production Function Parameters," *Economics Letters* 112 (August 2011), 205-206.

Christiano, Lawrence J., and Fitzgerald, Terry J., "The Band Pass Filter," *International Economic Review* 44 (May 2003), 435-465.

Cummins, Jason G., and Violante, Giovanni L., "Investment-Specific Technical Change in the United States (1947–2000): Measurement and Macroeconomic Consequences," *Review of Economic Dynamics* 5 (2002), 243-284.

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.

Engle, Robert, F., and Granger, Clive W.J., "Co-integration and Error Correction: Representation, Estimation, and Testing," *Econometrica* 55 (March 1987), 251-276.

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).

Gravelle, Jane G., "Historical Effective Marginal Tax Rates on Capital Income," Congressional Research Service Report for Congress (November 24, 2006).

Greenwood , Jeremy, Hercowitz, Zvi, and Krusell, Per, "Long-Run Implications of Investment-Specific Technological Change," *American Economic Review* 87 (June 1997), 342-362.

Griffin, James M., "Engineering and Econometric Interpretations of Energy-Capital Complementarity: Comment," *American Economic Review* 71 (December 1981), 1100-1104.

Hayashi, Fumio, *Econometrics*, (Princeton: Princeton University Press, 2000).

Herrendorf, Berthold, Herrington, Christopher, and Valentinyl, Akos, "Sectoral Technology and Structural Transformation," *American Economic Journal – Macroeconomics* 7 (October 2015), 104-133.

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

Hodrick, Robert J., and Prescott, Edward C., "Post-War U.S. Business Cycles: An Empirical Investigation," *Journal of Money, Credit and Banking* 29 (February 1997), 1-16.

Jones, Ronald W., and Ruffin, Roy J., "Trade and Wages: A Deeper Investigation," *Review of International Economics* 16 (May 2008), 234-249.

Jorgenson, Dale W., "Capital Theory and Investment Behavior," *American Economic Review* 53 (1963), 247-259.

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.

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, McAdam, Peter, and Willman, Alpo, "The Normalized CES Production Function: Theory and Empirics," *Journal of Economic Surveys* 26 (2012), 769-799.

Klump, Rainer, McAdam, Peter, and Willman, Alpo, "Factor Substitution and Factor-Augmenting Technical Progress in the United States: A Normalized Supply-Side System Approach," *The Review of Economics and Statistics* 89 (2007), 183-192.

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

Koopmans, Lambert H., *The Spectral Analysis of Time Series* (New York: Academic Press, 1974).

de La Grandville, Oliver, *Economic Growth – A Unified Approach* (Cambridge: Cambridge University Press, 2009).

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

León-Ledesma, Miguel, McAdam, Peter, and Willman, Alpo, "Production Technology Estimates and Balanced Growth," *Oxford Bulletin of Economics and Statistics* 77 (February 2015), 40-65.

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., "Two Illustrations of the Quantity Theory of Money," *American Economic Review* 70 (December 1980), 1005-1014.

Karabarbounis, Loukas, and Neiman, Brent, "The Global Decline of the Labor Share," *Quarterly Journal of Economics* 129 (February 2014), 61-103.

Mallick, Debdulal, "Substitution Elasticity and Balanced Growth," *Journal of Macroeconomics* 32 (December 2010), 1131–1142.

Mallick, Debdulal, "The Role of the Elasticity of Substitution in Economic Growth: A Cross-Country Investigation," *Labour Economics* 19 (October 2012), 682–694.

Newey, Whitney K., and West, Kenneth D. "A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," 55 *Econometrica* (1987), 703–708.

Oberfield, Ezra, and Raval, Devesh, "Micro Data and the Macro Technology," NBER Working Paper No. 20452 (September 2014).

Pesaran, M. Hashem, "Estimation and Inference in Large Heterogeneous Panels with a Multifactor Error Structure." *Econometrica* 74 (2006), 967–1012.

Pesaran, M. Hashem, "A Simple Panel Unit Root Test in the Presence of Cross Section Dependence," *Journal of Applied Econometrics* 22 (2007), 265-312.

Piketty, Thomas, *Capital in the Twenty-First Century* (Cambridge: Harvard University Press, 2014a).

Piketty, Thomas, "Technical Appendix for the Book *'Capital in the Twenty-First Century,"* (online at http://piketty.pse.ens.fr/files/capital21c/en/Piketty2014TechnicalAppendix.pdf , 2014b).

Piketty, Thomas, and Zucman, Gabriel, "Capital is Back: Wealth-Income Ratios in Rich Countries 1700-2010," *Quarterly Journal of Economics* 129 (August 2014), 1255-1310.

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.

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

Rognlie, Matthew, "A note on Piketty and diminishing returns to capital," MIT (June 2014).

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.

Summers, Lawrence H., "The Nonadjustment of Nominal Interest Rates: A Study of the Fisher Effect," in James Tobin (ed.), *Macroeconomics, Prices, and Quantities: Essays in Memory of Arthur Okun* (Washington: Brookings, 1983), 201-246.

Summers, Lawrence H., "The Inequality Puzzle," *Democracy: A Journal of Ideas* No. 32 (Spring 2014).

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

**TABLE 1: SUMMARY STATISTICS FOR FILTERED VARIABLES (*)
AND UNFILTERED VARIABLES**

Variable	Mean	Median	25th Quantile	75th Quantile	Within Industry Standard Deviation	Industry-Specific Time Variation
	(1)	(2)	(3)	(4)	(5)	(6)
A. Capital/Output						
$\Delta ky_{i,t}^{*}$	0.0108	0.0092	-0.0069	0.0282	0.0301	0.8724
$\Delta ky_{i,t}$	0.0109	0.0078	-0.0233	0.0434	0.0717	0.7746
$\Delta p_{i,t}^{KY*}$	-0.0058	-0.0050	-0.0350	0.0234	0.0504	0.8119
$\Delta p_{i,t}^{KY}$	-0.0053	-0.0003	-0.0731	0.0648	0.1427	0.8648
B. Capital/Labor						
$\Delta k\ell_{i,t}^{*}$	0.0298	0.0290	0.0084	0.0500	0.0386	0.8596
$\Delta k\ell_{i,t}$	0.0303	0.0296	-0.0136	0.0712	0.1001	0.8795
$\Delta p_{i,t}^{KL*}$	-0.0213	-0.0201	-0.0536	0.0097	0.0569	0.8249
$\Delta p_{i,t}^{KL}$	-0.0223	-0.0192	-0.0991	0.0528	0.1606	0.8757

Table notes are placed after the final table.

TABLE 2: OLS ESTIMATES OF EQUATION (5)
VARIOUS CRITICAL PERIODICITIES ($p^{\#}$) AND WINDOWS (q)

		q = 1	q = 3	q = 5
		(1)	(2)	(3)
$p^{\#} = 2$	σ	0.229	0.229	0.224
	(s.e.)	(0.022)	(0.023)	(0.025)
	[NW s.e.]	[0.020]	[0.022]	[0.023]
	{R^2}	{0.409}	{0.405}	{0.398}
$p^{\#} = 4$	σ	0.329	0.319	0.312
	(s.e.)	(0.016)	(0.017)	(0.017)
	[NW s.e.]	[0.022]	[0.024]	[0.026]
	{R^2}	{0.502}	{0.504}	{0.492}
$p^{\#} = 6$	σ	0.333	0.367	0.346
	(s.e.)	(0.016)	(0.016)	(0.017)
	[NW s.e.]	[0.022]	[0.027]	[0.026]
	{R^2}	{0.500}	{0.502}	{0.502}
$p^{\#} = 8$	σ	0.331	**0.406**	0.379
	(s.e.)	(0.016)	**(0.017)**	(0.017)
	[NW s.e.]	[0.022]	**[0.034]**	[0.032]
	{R^2}	{0.497}	**{0.503}**	{0.491}
$p^{\#} = 10$	σ	0.331	0.417	0.429
	(s.e.)	(0.016)	(0.018)	(0.019)
	[NW s.e.]	[0.022]	[0.037]	[0.040]
	{R^2}	{0.496}	{0.495}	{0.499}
$p^{\#} = 20$	σ	0.330	0.409	0.510
	(s.e.)	(0.016)	(0.018)	(0.020)
	[NW s.e.]	[0.022]	[0.037]	[0.049]
	{R^2}	{0.495}	{0.472}	{0.511}
$p^{\#} \to \infty$	σ	0.330	0.405	0.499
	(s.e.)	(0.016)	(0.018)	(0.020)
	[NW s.e.]	[0.022]	[0.037]	[0.047]
	{R^2}	{0.494}	{ 0.467}	{0.499}

Table notes are placed after the final table.

TABLE 3: OLS ESTIMATES OF EQUATION (5)
VARIETY OF DIFFERENT ESTIMATORS
VARIOUS CRITICAL PERIODICITIES ($p^{\#}$); FIXED WINDOW (q=3)

		Reordered Transformations	Cross-Sectional Dependence	Value-Added Output	Split-Sample 1960-1982	1983-2005	K/L Model
		(1)	(2)	(3)	(4)	(5)	(6)
$p^{\#} = 2$	σ	0.229	0.409	0.335	0.219	0.242	0.120
	[NW s.e.]	[0.022]	(0.115)	[0.028]	[0.035]	[0.036]	[0.050]
	{R^2}	{0.405}		{0.493}	{0.419}	{0.369}	{0.154}
$p^{\#} = 8$	σ	**0.395**	**0.391**	**0.543**	**0.336**	**0.460**	**0.333**
	[NW s.e.]	**[0.032]**	**(0.060)**	**[0.038]**	**[0.034]**	**[0.047]**	**[0.053]**
	{R^2}	**{0.500}**		**{0.652}**	**{0.472}**	**{0.496}**	**{0.339}**
$p^{\#} \to \infty$	σ	0.398	0.376	0.528	0.332	0.460	0.303
	[NW s.e.]	[0.035]	(0.069)	[0.037]	[0.036]	[0.052]	[0.046]
	{R^2}	{0.476}		{0.600}	{0.403}	{0.480}	{0.294}

Table notes are placed after the final table.

TABLE 4: IV AND OLS ESTIMATES OF EQUATION (5)
VARIETY OF DIFFERENT INSTRUMENTS
VARIOUS CRITICAL PERIODICITIES ($p^{\#}$); FIXED WINDOW (q=3)

		With Time Fixed Effects		Without Time Fixed Effects			
		Instruments					
		$p_{i,t-2}^{KY\&}$	$p_{i,t-2}^{K\&}$	Benchmark Model (OLS)	Corporate Tax Rate	Three Corporate Tax Rates	Three Corporate Tax Rates $p_{i,t-2}^{KY\&}$
		(1)	(2)	(3)	(4)	(5)	(6)
$p^{\#} = 2$	σ	0.359	0.316	0.254	0.440	0.513	0.456
	[NW s.e.]	[0.055]	[0.069]	[0.021]	[0.070]	[0.057]	[0.039]
	/F/	/7.14/	/6.63/		/29.49/	/15.74/	/24.64/
	(J p-value)					(0.001)	(0.002)
$p^{\#} = 8$	σ	**0.438**	**0.345**	**0.415**	**0.452**	**0.483**	**0.455**
	[NW s.e.]	**[0.062]**	**[0.087]**	**[0.029]**	**[0.103]**	**[0.080]**	**[0.050]**
	/F/	**/14.16/**	**/15.16/**		**/93.56/**	**/40.80/**	**/111.80/**
	(J p-value)					**(0.704)**	**(0.554)**
$p^{\#} \to \infty$	σ	0.464	0.355	0.414	0.460	0.493	0.459
	[NW s.e.]	[0.065]	[0.095]	[0.016]	[0.116]	[0.107]	[0.073]
	/F/	/13.77/	/13.69/		/92.09/	/34.40/	/57.37/
	(J p-value)					(0.088)	(0.141)

Table notes are placed after the final table.

TABLE 5: OLS ESTIMATES OF EQUATION (5)
HETEROGENEOUS INDUSTRY σ_i's AND AGGREGATED σ_{agg}
VARIOUS CRITICAL PERIODICITIES ($p^{\#}$); FIXED WINDOW (q=3)

Panel A: Heterogeneous Industry σ_i's

	Industry	p# = 8		p# = 2
		σ [NW s.e.]		σ [NW s.e.]
		(1)		(2)
1	Agriculture	0.289 [0.089]		0.105 [0.034]
2	Metal mining	0.667 [0.167]		0.113 [0.018]
3	Coal mining	0.305 [0.076]		0.215 [0.040]
4	Oil and gas extraction	0.649 [0.032]		0.517 [0.048]
5	Non-metallic mining	0.586 [0.029]		0.365 [0.029]
6	Construction	0.410 [0.098]		0.357 [0.038]
7	Food and kindred products	0.078 [0.034]		0.139 [0.020]
8	Tobacco	0.312 [0.140]		0.357 [0.042]
9	Textile mill products	0.204 [0.063]		0.094 [0.040]
10	Apparel	0.547 [0.098]		0.167 [0.043]
11	Lumber and wood	0.484 [0.067]		0.256 [0.028]
12	Furniture and fixtures	0.203 [0.081]		0.087 [0.051]
13	Paper and allied products	0.148 [0.054]		0.158 [0.056]
14	Printing, publishing and allied	0.484 [0.043]		0.126 [0.028]
15	Chemicals	0.210 [0.051]		0.359 [0.045]
16	Petroleum and coal products	0.294 [0.019]		0.155 [0.017]
17	Rubber and misc plastics	0.300 [0.059]		0.214 [0.021]
18	Leather	0.425 [0.059]		0.097 [0.061]
19	Stone, clay, glass	0.371 [0.037]		0.197 [0.024]
20	Primary metal	0.562 [0.045]		0.240 [0.036]
21	Fabricated metal	0.401 [0.034]		0.223 [0.024]
22	Machinery, non-electrical	0.483 [0.067]		0.292 [0.031]
23	Electrical machinery	0.486 [0.100]		0.193 [0.045]
24	Motor vehicles	0.365 [0.033]		0.270 [0.044]
25	Transportation equipment & ordnance	0.419 [0.052]		0.286 [0.091]
26	Instruments	0.570 [0.108]		0.440 [0.129]
27	Misc. manufacturing	0.246 [0.061]		0.144 [0.029]
28	Transportation	0.358 [0.078]		0.115 [0.036]
29	Communications	0.240 [0.112]		0.118 [0.098]
30	Electric utilities	0.231 [0.146]		0.101 [0.072]
31	Gas utilities	0.473 [0.080]		0.339 [0.042]
32	Trade	0.744 [0.054]		0.368 [0.037]
33	Finance, Insurance & Real Estate	1.160 [0.140]		0.085 [0.076]
34	Services	0.633 [0.094]		0.274 [0.073]
35	Government enterprises	0.272 [0.018]		0.238 [0.011]

TABLE 5 (CONTINUED): OLS ESTIMATES OF EQUATION (5)
HETEROGENEOUS INDUSTRY σ_i's AND AGGREGATED σ_{agg}
VARIOUS CRITICAL PERIODICITIES ($p^{\#}$); FIXED WINDOW (q=3)

Panel B: Aggregated σ_{agg} ($p^{\#}$=8; q=3)

	Heterogeneous				Homogeneous
	Capital Weights	Equal Weights	Capital Weights, Selective Industries		
	(1)	(2)	(3)		(4)
σ_{agg}	0.657	0.417	0.857		**0.406**
(s.e.)	(0.055)	(0.067)	(0.070)		**(0.017)**
[NW s.e.]	[0.043]	[0.072]	[0.089]		**[0.034]**

Panel C: Aggregated σ_{agg} ($p^{\#}$=2; q=3)

	Heterogeneous				Homogeneous
	Capital Weights	Equal Weights	Capital Weights, Selective Industries		
	(1)	(2)	(3)		(4)
σ_{agg}	0.196	0.223	0.177		**0.229**
(s.e.)	(0.052)	(0.074)	(0.082)		**(0.023)**
[NW s.e.]	[0.026]	[0.044]	[0.040]		**[0.022]**

TABLE NOTES:

Table 1: The table contains statistics for the variables entering the capital/output and capital/labor models in Panels A and B, respectively. The data are obtained from the webpage of Dale Jorgenson and represent inputs, output, and prices for 35 industries for the period 1960-2005; see fn. 11 for more details. In a given panel, the first and third rows are the filtered model variables; the second and fourth rows are comparable variables that have not been filtered by the LPF. The sample period is 1960-2005 for the starred/filtered series (first and third rows) and 1963-2002 for the unstarred/unfiltered series (second and fourth rows) before first differencing; the different sample periods ensure that the same data are used in computing the statistics for all entries. For a given industry, the effective time dimension equals these 46 data points less 2q for the construction of the LPF less one for first differencing. The statistics reported in this table are based on an LPF of $p^{\#} = 8$ and $q = 3$. The data underlying the moments in columns 1 to 5 have had industry means removed by first-differencing. Column 6 reports the variance due to industry-specific time variation (i.e., where the variance due to industry and common time effects have been removed) as a percentage of the overall variance in the series; this statistic is $(1 - R^2)$ from the following auxiliary regression for each variable appearing the table, $\Delta x_{i,t}$:

$\Delta x_{i,t} = a + b_t + r_{i,t}$, where a and b_t are estimated parameters and $r_{i,t}$ is a residual.

Table 2: OLS estimates of σ are based on equation (5) and a panel data for 35 industries. The sample period is 1960-2005 before applying the LPF and differencing. For a given industry, the effective time dimension equals these 46 data points less 2q for the construction of the LPF less one for first differencing. We use the maximum amount of available data, and hence the sample size changes with q. Figures in parentheses (s.e.) are White (1980) heteroskedasticity corrected standard errors. Figures in brackets [NW s.e.] are Newey-West (1994) heteroskedasticity-autocorrelation corrected (HAC) standard errors with (T-1) lags, which is equivalent to using a Bartlett kernel with a bandwidth of T. 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 based on the LPF parameters of $p^{\#} = 8$ and $q = 3$ and is highlighted in boldface.

Table 3: OLS estimates of σ based on equation (5). See the notes to Table 2 for further information. The complete set of results (paralleling those in Table 2) is presented in Appendix VI. Column 1 reorders the transformations to achieve stationarity [II and IV.A; entries in brackets indicate the relevant section of the manuscript]. Column 2 expands the benchmark model to allow for cross-sectional dependence in the error term [IV.B and Appendix V]; the standard errors are not Newey-West corrected because of computational issues with STATA related to non-linearity in the parameters, hence the use of (.). Column 3 replaces gross output by value-added output [IV.C]. Columns 4 and 5 present split-sample results [IV.D]. Column 6 replaces the capital/output ratio and its relative price in the benchmark model by the capital/labor ratio (K/L) and its relative price [IV.F].

TABLE NOTES (continued):

Table 4: IV estimates (except for the OLS estimates in column 3) of σ based on equation (5). See the notes to Table 2 for further information. The complete set of results (paralleling those in Table 2) is presented in Appendix VI. Figures between forward slashes are F-statistics computed from the first-stage auxiliary regression of $\Delta p_{i,t}^{KY*}$ on the instrument (and time dummies, but only for columns 1 and 2) that assess instrument relevance. The null hypothesis of a weak instrument is evaluated at the 5% level and rejected for F's greater than or equal to 8.96 (columns 1, 2, and 4), 12.83 (column 5), and 15.09 (column 6) and are taken from Stock, Wright, and Yogo (2002, Table 1). Figures in parentheses are p-values for the Hansen J-statistics and are presented only for the overidentified models in columns 5 and 6. Columns 1 [IV.E] and 2 [IV.F] use $p_{i,t-2}^{KY\&}$ and $p_{i,t-2}^{K\&}$ as instruments, respectively, and these instruments are constructed from the $p_{i,t}^{KY}$ and $p_{i,t}^{K}$ series with equation (6) as a one-sided filter (h = $\{0, -q\}$) and lagged two periods. Column 3 contains OLS estimates of the benchmark model without time fixed effects (i.e., $\tau_t^{KY} = 0$ in equation (5)). Columns 4, 5, and 6 [IV.F] do not include time fixed effects, and uses (tx_t^1), (tx_t^1, tx_t^2, tx_t^3), and ($tx_t^1, tx_t^2, tx_t^3, p_{i,t-2}^{KY\&}$) as instruments, respectively, where tx_t^n is an aggregate tax instrument from Gravelle (2006, Table 1; all quotations in this note are from Gravelle, 2006, p. 1). "These business tax rates reflect investments in equipment, structures, and inventory ." The tx_t^1 instrument is the corporate firm-specific tax rate**;** "if depreciation were allowed at economic rates and there were no subsidies, this rate would equal the corporate statutory tax rate." The tx_t^2 instrument is "the total rate on corporate investment, accounting for the deductibility of interest at the firm level and the taxation of interest, dividends, and capital gains at the individual level, as well as depreciation and subsidies." The tx_t^3 instrument is similar to tx_t^2 but for unincorporated business (proprietorships and partnerships).

Table 5: OLS estimates of industry σ_i's. See the notes to Table 2 for further information. Panel A contains estimates that allow the σ_i's to vary by industry. Panel B [V] contains aggregated σ_{agg}'s that are weighted averages as defined in equation (13) of the σ_i's in Panel A. See Appendix VII for a derivation of equation (13). Column 1 uses capital weights for all 35 industries; the weight for a given industry (w_i) is defined as follows, $w_i \equiv \sum_t K_{i,t} / \sum_t \sum_i K_{i,t}$. Column 2 uses equal weights for all 35 industries. Column 3 uses capital weights for six selective "post-industrial" industries: Construction (6), Transportation (28), Communications (29), Trade (32), Finance, Insurance and Real Estate (33) and Services (34). Column 4 contains the estimates for a homogeneous model with the restriction that $\sigma_i = \sigma \; \forall i$; that is, the benchmark estimates presented in Table 2. Standard errors are in parentheses or brackets and are computed with the following formula, $\sqrt{w'Vw}$, where w is a vector of weights and V is the variance-covariance matrix (either White (1980) or Newey-West (1994)) of the estimated coefficients.

Figure 1
Frequency Response Of α[ω : ω#, q]
Various Critical Periodicities (p#); q=3
See Section II For Details
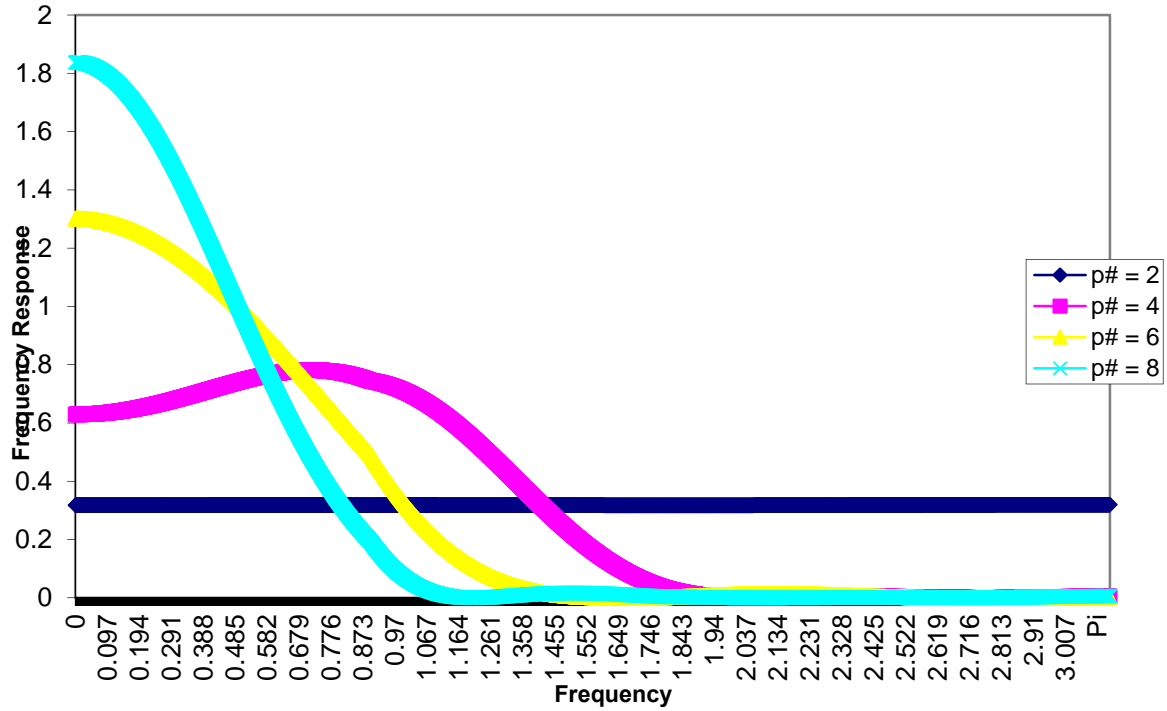
Figure 2
Frequency Response Of α[ω : ω#, q]
For Various Critical Periodicities (p#); q=3
See Section II For Details
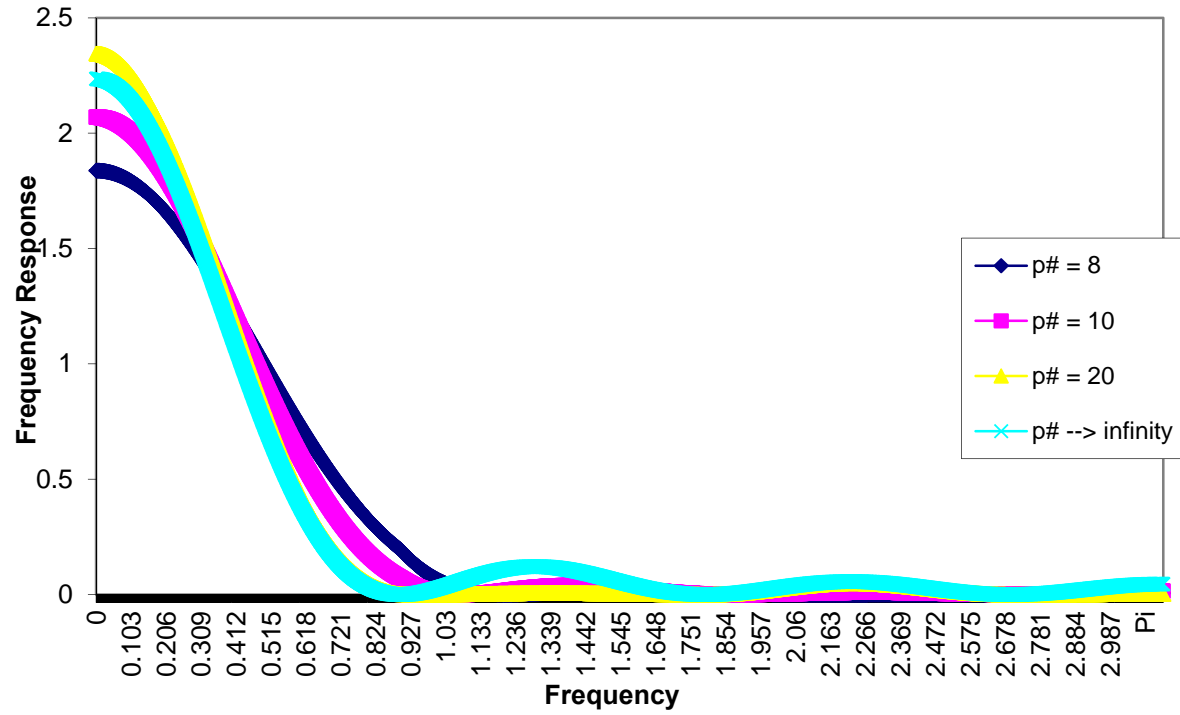
Figure 3
Frequency Response Of α[ω : ω#, q]
For The Ideal LPF And Various Windows (q); p# = 8
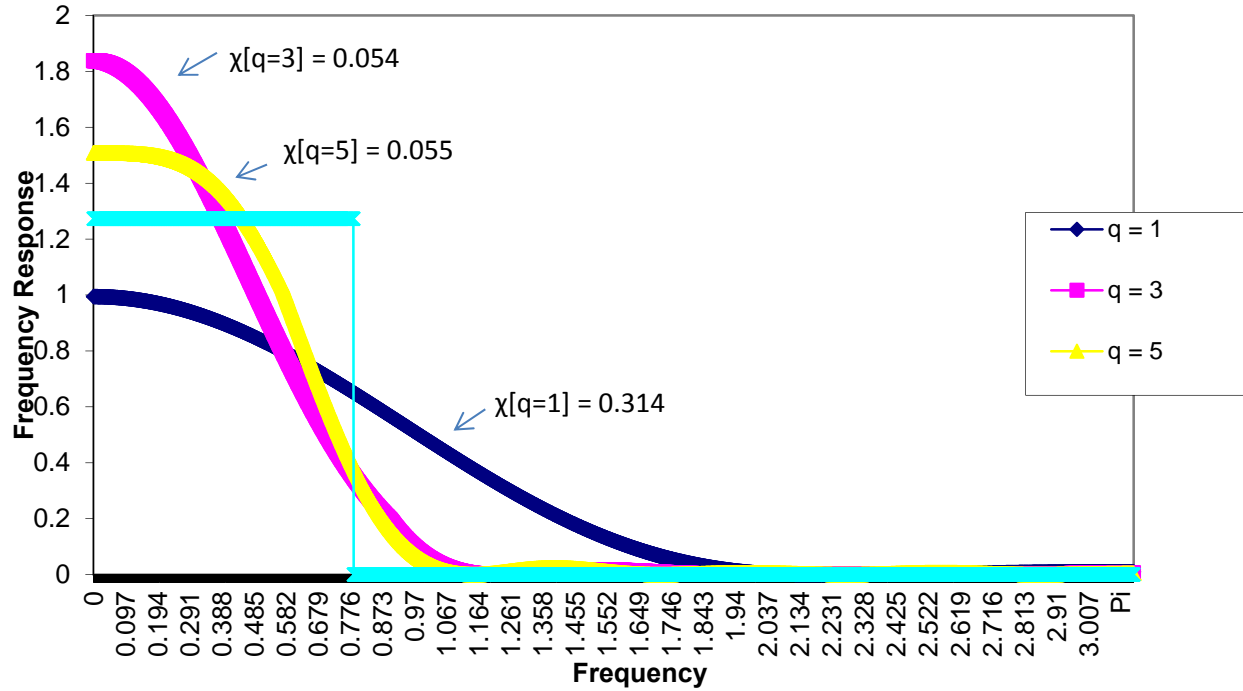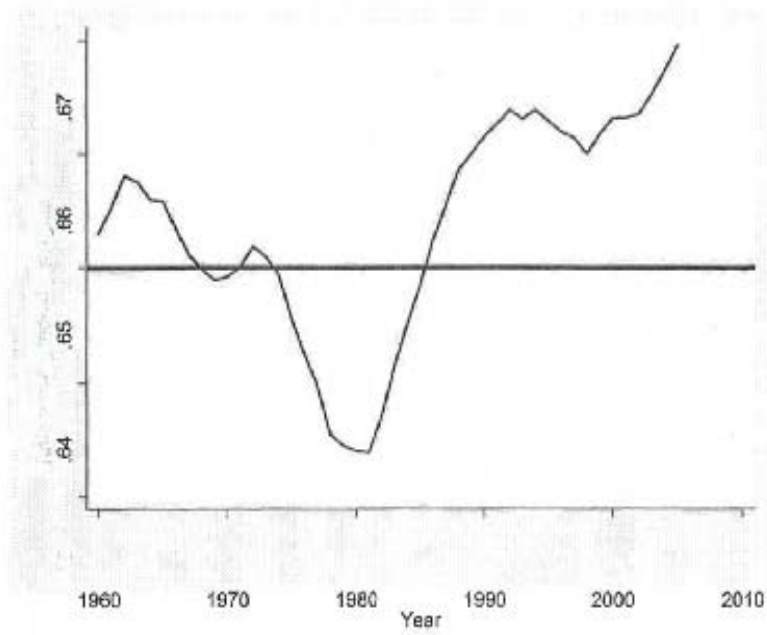**χ[q] Measures Contamination**
See Section II For Details

Figure 4
Time Plot Of $\breve{\sigma}_{agg,t}$; Equation (14)
Horizontal Line, σ_{agg}; Equation (13)

Not For Publication

Appendices:

The Substitution Elasticity, Factor Shares, Long-Run Growth, And

The Low-Frequency Panel Model

Appendix I:

A Crucial Assumption For Long-Run Economic Growth: The Value of σ

This appendix discusses the role played by σ in understanding a variety of issues concerning long-run economic growth – the possibility of perpetual growth or decline, the level of steady-state income per capita, the speed of convergence, the plausibility of the Solow growth model, and the role of technical change. It expands on the brief discussion found in the introductory section of the paper, and makes precise the definition of the "Solow Interval."

A. The Possibility Of Perpetual Growth Or Decline

The neoclassical revolution in growth theory introduced a variable capital/output ratio in place of the Harrod-Domar knife-edge and thus placed the burden of achieving equilibrium on the properties of the production function. When σ equals unity, the capital/labor ratio ($k \equiv K / L$) converges to a positive, finite value because, as k moves towards its limiting values of 0 (or ∞), the marginal product of capital and the average product of capital both tend to ∞ (or 0). With the Inada conditions satisfied, capital accumulation, as determined by Solow's fundamental equation of motion for k, converges to zero.

However, when σ departs from unity, the Inada conditions may fail, and perpetual decline or perpetual growth may occur. In their textbook on economic growth, Burmeister and Dobell (1970, p. 34) refer to situations where $\sigma \neq 1$ as "troublesome cases" because they do not yield balanced growth paths.[1] This appendix examines these "troublesome cases" and the role of

---

[1] 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 $\sigma = 1$ is a sufficient but not necessary condition for a balanced growth path. To treat σ as a free parameter determined by the theory runs dangerously close to the fallacy of affirming the consequent. An alternative approach -- represented in the models developed by Acemoglu (2003, 2009), Antrás (2004), Eicher and Turnovsky (1999), and Turnovsky and Smith (2006) -- interprets cases where

σ in generating non-standard growth behavior. Two conditions must hold: (1) $\sigma \neq 1$ and (2) certain relation between σ and the rates of saving (s), depreciation (δ), and population growth (n) and another production function parameter, ϕ, which is related to factor income shares.

We first consider the Solow's equation of motion for the capital/labor ratio ($k \equiv K/L$) and the limiting behavior of the marginal product of capital ($MPK[k:\sigma]$) and the average product of capital ($APK[k:\sigma]$).[2] The well-known equation of motion in the neoclassical growth model is as follows,

$$\dot{k}/k = s*APK[k:\sigma]-(n+\delta),$$ (I-1)

where s, n, and δ are the rates of saving, population growth, and depreciation, respectively. The $MPK[k:\sigma]$ and the $APK[k:\sigma]$ are derived from the following intensive form of the CES production function (where we have ignored in this appendix the role of technical change),

$$f[k:\sigma] = \{\phi k^{((\sigma-1)/\sigma)} + (1-\phi)\}^{(\sigma/(\sigma-1))} \qquad 0 < \phi < 1,$$ (I-2)

where $f[k:\sigma]$ is a per capita neoclassical production function depending on σ and ϕ (the capital distribution parameter). The $MPK[k:\sigma]$ and the $APK[k:\sigma]$ follow from equation (I-2) as follows,

$$MPK[k:\sigma] \equiv f_k[k:\sigma] = \phi\left\{\phi + (1-\phi)k^{((1-\sigma)/\sigma)}\right\}^{(1/(\sigma-1))},$$ (I-3)

$$APK[k:\sigma] \equiv \frac{f[k:\sigma]}{k} = \left\{\phi + (1-\phi)k^{((1-\sigma)/\sigma)}\right\}^{(\sigma/(\sigma-1))} = \left\{MPK[k:\sigma]/\phi\right\}^{\sigma}.$$ (I-4)

We are now in a position to assess the impacts of the two conditions. Condition (1) implies that the $MPK[k:\sigma]$ fails to satisfy one of the Inada conditions and that this positive, finite limit can affect the $APK[k:\sigma]$ such that no root exists for equation (I-1) for any value of

---

$\sigma \neq 1$ as quite interesting, suggesting needed modifications to the standard growth model and highlighting the key role played by σ.

[2] The analysis is this sub-section of Appendix I draws on the presentations of the neoclassical growth model in Barro and Sala-i-Martin (1995, Section 1.3.3), Klump and Preissler (2000), Klump and de La Grandville (2000), de La Grandville (1989, 2009), de La Grandville and Solow (2004, which also discusses how increases in σ expand production possibilities in a manner similar to exogenous technical progress), and especially Mallick (2010).

k. The limits for $MPK[k:\sigma]$ and $APK[k:\sigma]$ as $k \to \infty$ (equations (I-5)) and as $k \to 0$ (equations (I-6)) are as follows,

$k \to \infty$

$$\sigma > 1: \lim_{k \to \infty} MPK[k:\sigma] = \lim_{k \to \infty} APK[k:\sigma] = \phi^{(\sigma/(\sigma-1))} < 1, \qquad \text{(I-5a)}$$

$$\sigma = 1: \lim_{k \to \infty} MPK[k:\sigma] = \lim_{k \to \infty} APK[k:\sigma] = 0, \qquad \text{(I-5b)}$$

$$\sigma < 1: \lim_{k \to \infty} MPK[k:\sigma] = \lim_{k \to \infty} APK[k:\sigma] = 0. \qquad \text{(I-5c)}$$

$k \to 0$

$$\sigma > 1: \lim_{k \to 0} MPK[k:\sigma] = \lim_{k \to 0} APK[k:\sigma] = \infty, \qquad \text{(I-6a)}$$

$$\sigma = 1: \lim_{k \to 0} MPK[k:\sigma] = \lim_{k \to 0} APK[k:\sigma] = \infty, \qquad \text{(I-6b)}$$

$$\sigma < 1: \lim_{k \to 0} MPK[k:\sigma] = \lim_{k \to 0} APK[k:\sigma] = \phi^{(\sigma/(\sigma-1))} > 1. \qquad \text{(I-6c)}$$

The Inada conditions are met when $\sigma = 1$ (equations (I-5b) and (I-6b)) or, in one direction, when $\sigma > 1$ (equation (I-6a)) or when $\sigma < 1$ (equation (I-5c)). However, they fail in one direction when $\sigma > 1$ (equation (I-5a)) or when $\sigma < 1$ (equation (I-6c)).

The implications of these two failures are that perpetual growth or decline is possible. These non-standard cases depend on a second necessary condition comparing these limiting values to the critical values of the $APK[k:\sigma]$ that set $\dot{k}/k$ equal to zero in equation (I-1). The critical value of $APK[k:\sigma]$ can be stated in terms of the associated critical value of σ, σ_C, by setting equation (I-1) to zero and solving for σ_C, which will depend on four other model parameters (s. n, δ, and ϕ),

$$\sigma_C \equiv g[s, n, \delta, \phi] = \frac{\log[s/(n+\delta)]}{\log[(s\phi)/(n+\delta)]} \gtreqless 1 . \qquad \text{(I-7)}$$

Whether σ_C is above, equal to, or below 1.0 depends on the values of the four parameters.

To assess the possibility of perpetual growth, perpetual decline, or balanced growth, we will consider only variations in the saving rate, holding the other three parameters fixed. Three cases need to be considered:

- **Case 1, Perpetual Growth:** If the saving rate is sufficiently large so that $s > (n + \delta) / \phi$, then, since $0 < \phi < 1$, the numerator is greater than the denominator, and equation (I-7) defines the critical value defining perpetual growth, $\sigma_{PG} > 1$. When $\sigma > \sigma_{PG} > 1$, no increase in capital accumulation, however large, will be able to push the $MPK[k : \sigma]$ and $APK[k : \sigma]$ sufficiently low to terminate perpetual growth.

- **Case 2, Perpetual Decline:** Alternatively, if the saving rate is sufficiently low so that $s < (n + \delta)$, then, since $0 < \phi < 1$, the numerator is less than the denominator, and equation (I-7) defines the critical value defining perpetual decline, $\sigma_{PD} < 1$. When $\sigma < \sigma_{PD} < 1$, no decrease in capital accumulation (bounded below by zero) will be able to push the $MPK[k : \sigma]$ and $APK[k : \sigma]$ sufficiently high to escape perpetual decline.

- **Case 3, Balanced Growth:** If the saving rate takes on intermediate values relative to Cases 1 and 2 ($(n + \delta) / \phi \geq s \geq (n + \delta)$), then the Inada conditions are satisfied, and balanced growth is achieved. We refer to this range of σ's as the Solow Interval.

The critical values depend on four parameters -- s, n, δ, and ϕ. We assume the following values for three of the parameters,

- $n = 0.01$,
- $\delta = 0.10$ (a standard value for business plant and equipment),
- $\phi = 0.33$ (this production function parameter is not identified in the first-order conditions estimated in this paper. We assume that it equals the capital's income share).

The most interesting variation is with the saving rate, and we evaluate the critical value with alternative values of s in Appendix Table I,

APPENDIX TABLE I: CRITICAL VALUES OF σ (σ_C)
FOR ALTERNATIVE SAVING RATES

Saving Rate	σ_C		Saving Rate	σ_C
0.01	0.68		0.14	-0.28
0.02	0.61		0.15	-0.39
0.03	0.54		0.16	-0.51
0.04	0.48		0.17	-0.65
0.05	0.42		**0.18**	**-0.80**
0.06	0.35		0.19	-0.97
0.07	0.29		0.20	-1.17
0.08	0.22		0.21	-1.40
0.09	0.15		0.22	-1.67
0.10	0.08		0.23	-1.99
0.11	0.00		0.24	-2.37
0.12	-0.09		0.25	-2.85
0.13	-0.18			

For s = 0.01, the critical value for σ is 0.68, and it declines monotonically for the saving rates considered in Appendix Table I-1. (There is a singularity (not shown in the table) for $0.33 < s < 0.34$. For $s \geq 0.35$, σ_C is positive, declines monotonically in s and, as $s \to 1$, $\sigma_C \to 2.01$.) Over the last twenty years, the average saving rate is 0.18 (the ratio of gross saving (NIPA Table 5.1, Line 1) to GDP (NIPA Table 1.1.5, Line 1) averaged for the period 1993 to 2012). For this value of s, $\sigma_C < 0$. Thus, it is impossible for the economy to enter a state of perpetual decline (since σ cannot be less than zero), and the economy exhibits balanced growth.

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, 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).[3] De La Grandville (1989, 2009) argues theoretically that a relative price change (e.g., a decrease in the price of capital) leads to relatively more output the higher

---

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

the value of σ. (He also identifies a second channel depending on σ -- a higher substitution elasticity permits a greater flow of resources between sectors with different factor intensities.) Yuhn (1991) and Mallick (2012) find empirical support of this hypothesis for South Korea and across 90 countries, respectively.

C. *The Speed of Convergence*

The speed of convergence to the steady-state depends on σ through capital accumulation. In a calibrated neoclassical growth model, Turnovsky (2002, pp. 1776-1777) finds 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 increases to 1.0, 1.2, and 1.5, respectively.[4] Klump and Preissler (2000, equation 14) show that the speed of convergence is sensitive to the relation between the initial and steady-state capital intensities.

D. *The Plausibility Of The Solow Growth Model*

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 can be implausibly high when some part of growth is due to transitional dynamics. When transitional dynamics are important, R increases dramatically with σ. Mankiw (1995, p. 287, equation (13)) also investigates the relation between σ and R in terms of the following equation,

$$R = \left(y\right)^{-\theta} Q \qquad \theta \equiv \left(1/\sigma\right)*\left(\left(1-\mu^{K}\right)/\mu^{K}\right), \tag{I-8}$$

where μ^{K} is capital's factor share, y is per capita income, and Q represents other factors affecting the return to capital. Equation (I-8) links rates of return to income and can be used to evaluate rate of return differentials between poor and rich countries. For example, if $\sigma = 4.0$,

---

[4] 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.

the differential is only about 3 percent.[5] But if σ falls to 1.0 or 0.5, the differential becomes implausibly large, increasing to 100 and 10,000 respectively, clearly raising concerns about the standard growth model.

E. Role Of Technical Change

The role played by technical change in growth models depends on the value of σ. If $\sigma = 1$, factor-augmenting technical change becomes indistinguishable from neutral technical change. If $\sigma < 1$ (>1) factor-augmenting technical change lowers (raises) the intensity of the usage of that factor (see equation (II-5) in Appendix II). Acemoglu (2003) examines the relations among technical change, 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 what percentage 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, and technical efficiency plays little role. The percentage decreases in σ and drops to 40% for $\sigma = 1.0$ (the Cobb-Douglas case) and 25% for $\sigma = 1.5$. Caselli (2005, p. 737) concludes "that the Cobb-Douglas assumption is a very sensitive one for development accounting."

---

[5] 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.

Appendix II:

Specifying The Marginal Product Of Capital

With Neutral And Factor-Augmenting Technical Change

This appendix presents the details of the derivation of the marginal product of capital when there is both neutral and factor-augmenting technical change, and the derivation shows that the latter has no impact on the specification of the estimating equation used in this study. We assume that production possibilities are described by the following CES technology that relates output ($Y_{i,t}^*$) to capital ($K_{i,t}^*$), labor ($L_{i,t}^*$), and neutral technical progress ($A_{i,t}^*$), and factor-augmenting technical progress on capital and labor ($A_{i,t}^{K*}$ and $A_{i,t}^{L*}$, respectively) for industry i at time t,

$$Y_{i,t}^* = Y[K_{i,t}^*, L_{i,t}^*, A_{i,t}^*, A_{i,t}^{K*}, A_{i,t}^{L*}], \tag{II-1}$$

$$= A_{i,t}^* \left\{ \phi(A_{i,t}^{K*} K_{i,t}^*)^{[(\sigma-1)/\sigma]} + (1-\phi)(A_{i,t}^{L*} L_{i,t}^*)^{[(\sigma-1)/\sigma]} \right\}^{[\sigma/(\sigma-1)]}$$

where ϕ is the capital distribution parameter and σ is the elasticity of substitution between labor and capital. We use the value-added form of the CES production function for convenience.

The derivative of $Y_{i,t}^*$ with respect to $K_{i,t}^*$, $Y'_{i,t}^*$, is computed from equation (II-1) as follows,

$$Y'_{i,t}^* = [\sigma/(\sigma-1)]A_{i,t}^* \left\{ \phi(A_{i,t}^{K*} K_{i,t}^*)^{[(\sigma-1)/\sigma]} + (1-\phi)(A_{i,t}^{L*} L_{i,t}^*)^{[(\sigma-1)/\sigma]} \right\}^{[[\sigma/(\sigma-1)]-1]}$$

$$* [(\sigma-1)/\sigma]\phi(A_{i,t}^{K*} K_{i,t}^*)^{[[(\sigma-1)/\sigma]-1]} A_{i,t}^{K*}. \tag{II-2}$$

Since $((\sigma-1)/\sigma)-1 = -1/\sigma$, we can rewrite equation (II-2) as follows,

$$Y'_{i,t}^* = \phi \, K_{i,t}^{*[-1/\sigma]}$$

$$A_{i,t}^* \left\{ \phi(A_{i,t}^{K*} K_{i,t}^*)^{[(\sigma-1)/\sigma]} + (1-\phi)(A_{i,t}^{L*} L_{i,t}^*)^{[(\sigma-1)/\sigma]} \right\}^{[\sigma/(\sigma-1)]}$$

$$\left\{ \phi(A_{i,t}^{K*} K_{i,t}^*)^{[(\sigma-1)/\sigma]} + (1-\phi)(A_{i,t}^{L*} L_{i,t}^*)^{[(\sigma-1)/\sigma]} \right\}^{-1} \tag{II-3}$$

$$A_{i,t}^{K*[(\sigma-1)/\sigma]}.$$

In equation (II-3), the second line equals $Y_{i,t}^*$ per equation (II-1), and the third line equals the product of $Y_{i,t}^*$ and $A_{i,t}^*$ raised to the appropriate powers,

$$Y{'}_{i,t}^* = \phi \; K_{i,t}^{*\,[-1/\sigma]}$$
$$\frac{Y_{i,t}^*}{Y_{i,t}^{*[(1-\sigma)/\sigma]} \; A_{i,t}^{*[(\sigma-1)/\sigma]}} \quad , \tag{II-4}$$
$$A_{i,t}^{K*[(\sigma-1)/\sigma]}$$

which can be rewritten as follows,

$$Y{'}_{i,t}^* = \phi \; K_{i,t}^{*\,[-1/\sigma]} \; Y_{i,t}^{*[1/\sigma]} \; A_{i,t}^{*[(\sigma-1)/\sigma]} \; A_{i,t}^{K*[(\sigma-1)/\sigma]} \quad . \tag{II-5a}$$

$$= \phi \; ((Y_{i,t}/K_{i,t})^*)^{[1/\sigma]} \; U_{i,t}^{KY\,[1/\sigma]} , \tag{II-5b}$$

$$U_{i,t}^{KY\,[1/\sigma]} \equiv A_{i,t}^{*[\sigma-1]} \; A_{i,t}^{K*[\sigma-1]}. \tag{II-5c}$$

A profit-maximizing firm will equate the marginal product of capital in equation (II-5a) to the user cost of capital, the price of capital divided by the price of output,

$$(P_{i,t}^K / P_{i,t}^Y)^* = \phi \; ((Y_{i,t}/K_{i,t})^*)^{[1/\sigma]} \; U_{i,t}^{KY\,[1/\sigma]}. \tag{II-6}$$

Equation (II-6) can be rearranged to isolate the capital/output ratio on the left-side,

$$(K_{i,t}/Y_{i,t})^* = \phi^\sigma \; ((P_{i,t}^K / P_{i,t}^Y)^*)^{-\sigma} \; U_{i,t}^{KY} , \tag{II-7}$$

which is equation (2) in the text except for the inclusion of $A_{i,t}^{K*}$ in the error term (cf. equation (II-5c). Thus, factor-augmenting technical change has no impact on the specification of the estimating equation for σ, and the only implication is a warning about possible correlation between the error term and the regressor.

Appendix III:

Data Transformations And The Frequency Response Scalars

Our estimation strategy is designed to emphasize long-run variation, and Section II uses spectral analysis to evaluate our approach and the choices of $\omega^{\#}$ and q. This appendix provides some analytic details underlying the results stated and used in Section II.

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

$$x^*_{i,t}[\omega^{\#}, q] = \sum_{h=-q}^{q} d_h[\omega^{\#}] \, x_{i,t-h}, \tag{III-1a}$$

$$y^*_{i,t}[\omega^{\#}, q] = \ln[x^*_{i,t}[\omega^{\#}, q]], \tag{III-1b}$$

$$z^*_{i,t}[\omega^{\#}, q] = \Delta y^*_{i,t}[\omega^{\#}, q], \tag{III-1c}$$

where $x_{i,t}$ represents the raw data series, either $(K_{i,t} / Y_{i,t})$ or $p^{KY}_{i,t}$. The spectra corresponding to the $x^*_{i,t}[.]$, $y^*_{i,t}[.]$, and $z^*_{i,t}[.]$ output series in equations (III-1) are defined over the interval $\omega = [0, \pi]$ as the product of the spectrum for an input series and a scalar that is nonnegative, real, and may be depend on ω, $\omega^{\#}$, or q,

$$g_{x*}[\omega, \omega^{\#}, q] = \alpha[\omega, \omega^{\#}, q] \, g_x[\omega], \tag{III-2a}$$

$$g_{y*}[\omega, \omega^{\#}, q] = \beta \, g_{x*}[\omega], \tag{III-2b}$$

$$g_{z*}[\omega, \omega^{\#}, q] = \gamma[\omega] \, g_{y*}[\omega], \tag{III-2c}$$

where $g_x[\omega]$ is the spectrum for the raw series and the scalars are defined as follows,

$$\alpha[\omega, \omega^{\#}, q] = a[\omega^{\#}, q] \left\{ \begin{array}{l} (\omega^{\#} / \pi) + 2 \sum_{h=1}^{q} \cos[h\,\omega] \, d'_h[\omega^{\#}] \\ + \theta[\omega^{\#}, q] \, \{(1 - \cos[\omega(2q+1)]) / (1 - \cos[\omega])\}^{1/2} \end{array} \right\}^2 \tag{III-3a}$$

$$\beta = b \, (\mu_{x*})^{-2}, \tag{III-3b}$$

$$\gamma[\omega] = c \, 2 \, (1 - \cos[\omega]), \tag{III-3c}$$

where μ_{x*} equals the unconditional expectation of $x^*_{i,t}[.]$. To ensure comparability in the analyses to follow that vary $\omega^\#$ and q, the areas under the spectra from 0 to π are normalized to one by an appropriate choice of normalizing constants, $a[\omega^\#,q]$, b, and c in equations (III-3).

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

$$\alpha[\omega,\omega^\#,q] = a[\omega^\#,q]\left\{\sum_{h=-q}^{q} e^{-ih\omega}d_h[\omega^\#]\right\}\left\{\sum_{h=-q}^{q} e^{ih\omega}d_h[\omega^\#]\right\}. \tag{III-4}$$

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 nearly identical. The $d_h[.]$'s appearing in the summations are separated into $\theta[.]$ and the $d'_h[.]$'s (cf. equations (6)). 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 (III-4) can be written as follows,

$$\alpha[\omega,\omega^\#,q] = a[\omega^\#,q]\left\{\theta[\omega^\#,q]\sum_{h=-q}^{q} e^{ih\omega} + (\pi\omega) + \sum_{h=1}^{q}(e^{-ih\omega}+e^{ih\omega})d'_h[\omega^\#]\right\}^2. \tag{III-5}$$

The first sum of exponential terms in equation (III-5) is evaluated based on Sargent (1987, p. 275),

$$\sum_{h=-q}^{q} e^{ih\omega} = \left\{\left(\sum_{h=-q}^{q} e^{ih\omega}\right)^2\right\}^{1/2} = \left\{((1-\cos[(2q+1)\omega]/(1-\cos[\omega])\right\}^{1/2}. \tag{III-6}$$

The second sum of exponential terms in equation (III-5) is evaluated with the Euler relations, $e^{\pm ih\omega} = \cos[h\omega] \pm i\,\sin[h\omega]$,

$$\sum_{h=1}^{q}(e^{-ih\omega}+e^{ih\omega})d'_h[\omega^\#] = 2\sum_{h=1}^{q}\cos[h\omega]\,d'_h[\omega^\#]. \tag{III-7}$$

The β 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 ($x^*_{i,t}[.]$).

The $\gamma[\omega]$ scalar is based on the well-known formula for the first-difference transformation (Hamilton, 1994, equation 6.4.8).

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,

$$g_{z*}[\omega,\omega^{\#},q] = \alpha[\omega,\omega^{\#},q] \ \beta \ \gamma[\omega] \ g_x[\omega]. \tag{III-8}$$

Equation (III-8) allows us to examine the extent to which our estimation strategy emphasizes long-run frequencies. Since the spectra for the raw series ($g_x[\omega]$) and the scalars associated with the logarithmic and first-difference transformations (β and $\gamma[\omega]$, respectively) do not depend on $\omega^{\#}$ or q, their impacts on the data will be absorbed in the normalizing constants ($a[\omega^{\#},q]$, b, and c), and hence they will not affect relative comparisons. Alternative values of $\omega^{\#}$ or q, will only affect the $LPF[\omega^{\#},q]$ and the associated scalar, $\alpha[\omega,\omega^{\#},q]$.

Appendix IV:
Stationarity Properties Of The Model Variables

To assess stationarity of the variables used in the regression models, we use the panel unit root test proposed by Pesaran (2007) that extends the standard augmented Dickey-Fuller test to allow for cross-sectional dependence in panel data. For a given variable $X_{i,t} = \left\{ (K_{i,t} / Y_{i,t}),\ (P_{i,t}^{K} / P_{i,t}^{Y}),\ \Delta Ln(K_{i,t} / Y_{i,t}),\ \Delta Ln(P_{i,t}^{K} / P_{i,t}^{Y}), \right\}$, we estimate the following auxiliary equation,

$$\Delta X_{i,t} = a_i + b_i X_{i,t-1} + \bar{b}_i \bar{X}_{t-1} + \sum_{j=1}^{J'} d_{i,j} \Delta X_{i,t-j} + \sum_{j=0}^{J'} \bar{d}_{i,j} \overline{\Delta X}_{t-j} + g_i\, t + u_{i,t}, \quad \text{(IV-1a)}$$

$$u_{i,t} = r_i\, u_{i,t-1} + \varepsilon_{i,t}, \qquad\qquad\qquad \text{(IV-1b)}$$

$$\mu_b = \sum_{i=1}^{35} b_i / 35, \qquad\qquad\qquad \text{(IV-1c)}$$

where $\bar{X}_t$ is a cross-section average of $X_{i,t}$, t is a time index, and the remaining lower case roman letters are parameters to be estimated. The lag length (J') for the lagged dependent variable ($\Delta X_{i,t-j}$) and lagged difference in cross-section averages ($\overline{\Delta X}_{t-j}$) is determined by the need to absorb any serial correlation in the errors. The null hypothesis that $X_{i,t}$ has a unit root is evaluated by μ_b, the average of the estimated b_i coefficients. The critical values for μ_b are provided in Pesaran's Tables II.b and II.c for tests without and with a time trend ($g_i\, t$ in the above equation), respectively; they are reported in columns (3) to (5) in Appendix Table IV below.

The test statistics for the raw series are presented in panel A of Appendix Table IV, and they are negative and generally greater than the (negative) critical values, with the exception of the relative price terms in models with a deterministic trend, which is close to the 10% critical value. These results suggest that the two model variables are not stationary.

As discussed in sub-section IV.A. , the LPF is not strictly valid when applied to nonstationary data, and we exploit the commutative property of the filters (equation (7d)). In this case, we take logs and first-difference the data before applying the LPF. As shown in panel B of Appendix Table IV, the unit root null hypothesis is rejected for all four model variables with or without a deterministic trend. Thus, it is appropriate to apply the LPF to the logarithmically differenced data.

APPENDIX TABLE IV: TEST STATISTICS FOR STATIONARITY OF THE MODEL VARIABLES μ_b IN EQUATION (IV-1c)

Panel A: Raw Series

Variable	With Deterministic Trend $(g_i \neq 0)$	Critical Values			Without Deterministic Trend $(g_i = 0)$	Critical Values		
		10%	5%	1%		10%	5%	1%
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$K_{i,t} / Y_{i,t}$	-2.494	-2.550	-2.600	-2.720	-1.725	-2.050	-2.110	-2.230
$P^K_{i,t} / P^Y_{i,t}$	-2.852	-2.550	-2.600	-2.720	-2.085	-2.050	-2.110	-2.230
$K_{i,t} / L_{i,t}$	-1.856	-2.550	-2.600	-2.720	-1.722	-2.050	-2.110	-2.230
$P^K_{i,t} / P^L_{i,t}$	-2.730	-2.550	-2.600	-2.720	-2.448	-2.050	-2.110	-2.230

Panel B: Transformed Series; First-Differences Of Logs Of Raw Series

Variable	With Deterministic Trend $(g_i \neq 0)$	Critical Values			Without Deterministic Trend $(g_i = 0)$	Critical Values		
		10%	5%	1%		10%	5%	1%
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\Delta Ln[K_{i,t} / Y_{i,t}]$	-5.342	-2.550	-2.600	-2.720	-5.118	-2.050	-2.110	-2.230
$\Delta Ln[P^K_{i,t} / P^Y_{i,t}]$	-6.133	-2.550	-2.600	-2.720	-5.974	-2.050	-2.110	-2.230
$\Delta Ln[K_{i,t} / L_{i,t}]$	-5.702	-2.550	-2.600	-2.720	-5.495	-2.050	-2.110	-2.230
$\Delta Ln[P^K_{i,t} / P^L_{i,t}]$	-6.243	-2.550	-2.600	-2.720	-6.059	-2.050	-2.110	-2.230

Appendix V:

The Common Correlated Effects (CCE) Estimator

Cross-sectional dependence may not be fully captured by time fixed effects and hence becomes part of the error term. Correlation between these shocks and the regressors would lead to inconsistent estimates of σ. Even absent such correlation, the shocks will lead to biased standard errors. The common correlated effects (CCE) estimator introduced by Pesaran (2006) is feasible for panels with a large number of cross-section units (unlike the Seemingly Unrelated Regression framework), and it accounts for the effects of cross-sectional dependence by including cross-section averages (CSA's) of the dependent and independent variables as additional right-hand side variables,

$$\Delta ky_{i,t}^* = \zeta^{KY} - \sigma \Delta p_{i,t}^{KY*} + e_{i,t}^{KY} + \gamma_i \left(CSA\left[\Delta ky_{i,t}^* \right] - \sigma\, CSA\left[\Delta p_{i,t}^{KY*} \right] \right), \qquad \text{(V-I)}$$

where $CSA[.]$ is the cross-section average operator and the γ_i's are 35 additional parameters to be estimated. If the γ_i's in equation (V-I) are constrained to be 1.0 for all i, the specification would be equivalent to transforming the data by demeaning each variable with respect to its CSA, the standard way of controlling for time fixed effects with the least squares dummy variables (LSDV) estimator.

In general, the CSA's in the CCE estimator are formed with a set of state weights, v_j for $j = 1,\dots,J$, such that,

$$\bar{x}_t \equiv \sum_{j=1}^{J} v_j x_{j,t}, \qquad \sum_{j=1}^{J} v_j = 1. \qquad \text{(V-II)}$$

As shown by Pesaran (2006), the asymptotic properties of the CCE estimator are invariant to the choice of the v_j weights. The empirical work reported here is based on equal weighting ($v_j = 1/J$ for all j).

Appendix VI:

Complete Tables 3 And 4

TABLE 3: OLS ESTIMATES OF EQUATION (5)
VARIETY OF DIFFERENT ESTIMATORS
VARIOUS CRITICAL PERIODICITIES ($p^{\#}$); FIXED WINDOW (q=3)

		Reordered Transformations	Cross-Sectional Dependence	Value-Added Output	Split-Sample 1960-1982	1983-2005	K/L Model
		(1)	(2)	(3)	(4)	(5)	(6)
$p^{\#} = 2$	σ	0.229	0.409	0.335	0.219	0.242	0.120
	[NW s.e.]	[0.022]	(0.115)	[0.028]	[0.035]	[0.036]	[0.050]
	{R^2}	{0.405}		{0.493}	{0.419}	{0.369}	{0.154}
$p^{\#} = 4$	σ	0.323	0.375	0.442	0.291	0.350	0.262
	[NW s.e.]	[0.024]	(0.124)	[0.031]	[0.030]	[0.033]	[0.051]
	{R^2}	{0.514}		{0.611}	{0.539}	{0.446}	{0.274}
$p^{\#} = 6$	σ	0.361	0.370	0.503	0.316	0.413	0.304
	[NW s.e.]	[0.026]	(0.119)	[0.035]	[0.029]	[0.040]	[0.052]
	{R^2}	{0.503}		{0.641}	{0.510}	{0.472}	{0.315}
$p^{\#} = 8$	**σ**	**0.395**	**0.391**	**0.543**	**0.336**	**0.460**	**0.333**
	[NW s.e.]	**[0.032]**	**(0.060)**	**[0.038]**	**[0.034]**	**[0.047]**	**[0.053]**
	{R^2}	**{0.500}**		**{0.652}**	**{0.472}**	**{0.496}**	**{0.339}**
$p^{\#} = 10$	σ	0.406	0.291	0.550	0.341	0.472	0.333
	[NW s.e.]	[0.035]	(0.171)	[0.039]	[0.037]	[0.051]	[0.051]
	{R^2}	{0.494}		{0.641}	{0.442}	{0.499}	{0.333}
$p^{\#} = 20$	σ	0.401	0.415	0.534	0.335	0.464	0.310
	[NW s.e.]	[0.035]	(0.095)	[0.038]	[0.036]	[0.052]	[0.046]
	{R^2}	{0.479}		{0.607}	{0.407}	{0.484}	{0.301}
$p^{\#} \to \infty$	σ	0.398	0.376	0.528	0.332	0.460	0.303
	[NW s.e.]	[0.035]	(0.069)	[0.037]	[0.036]	[0.052]	[0.046]
	{R^2}	{0.476}		{0.600}	{0.403}	{0.480}	{0.294}

Table notes are placed after the final table in the manuscript.

TABLE 4: IV AND OLS ESTIMATES OF EQUATION (5)
VARIETY OF DIFFERENT INSTRUMENTS
VARIOUS CRITICAL PERIODICITIES ($p^{\#}$); FIXED WINDOW (q=3)

		With Time Fixed Effects		Without Time Fixed Effects			
		Instruments					
		$p_{i,t-2}^{KY\&}$	$p_{i,t-2}^{K\&}$	Benchmark Model (OLS)	Corporate Tax Rate	Three Corporate Tax Rates	Three Corporate Tax Rates $p_{i,t-2}^{KY\&}$
		(1)	(2)	(3)	(4)	(5)	(6)
$p^{\#}=2$	σ [NW s.e.] /F/ (J p-value)	0.359 [0.055] /7.14/	0.316 [0.069] /6.63/	0.254 [0.021]	0.440 [0.070] /29.49/	0.513 [0.057] /15.74/ (0.001)	0.456 [0.039] /24.64/ (0.002)
$p^{\#}=4$	σ [NW s.e.] /F/ (J p-value)	0.367 [0.049] /10.57/	0.314 [0.062] /11.93/	0.353 [0.023]	0.433 [0.085] /43.35/	0.478 [0.053] /28.09/ (0.078)	0.417 [0.037] /63.68/ (0.082)
$p^{\#}=6$	σ [NW s.e.] /F/ (J p-value)	0.412 [0.058] /12.99/	0.333 [0.077] /13.94/	0.387 [0.024]	0.445 [0.095] /70.56/	0.481 [0.064] /36.34/ (0.381)	0.440 [0.045] /88.33/ (0.244)
$p^{\#}=8$	**σ [NW s.e.] /F/ (J p-value)**	**0.438 [0.062] /14.16/**	**0.345 [0.087] /15.16/**	**0.415 [0.029]**	**0.452 [0.103] /93.56/**	**0.483 [0.080] /40.80/ (0.704)**	**0.455 [0.050] /111.80/ (0.554)**
$p^{\#}=10$	σ [NW s.e.] /F/ (J p-value)	0.450 [0.063] /14.30/	0.349 [0.091] /15.12/	0.423 [0.031]	0.456 [0.108] /99.65/	0.485 [0.091] /40.26/ (0.455)	0.462 [0.053] /117.58/ (0.594)
$p^{\#}=20$	σ [NW s.e.] /F/ (J p-value)	0.462 [0.065] /13.90/	0.354 [0.094] /14.00/	0.417 [0.031]	0.460 [0.115] /94.40/	0.492 [0.104] /35.57/ (0.120)	0.472 [0.055] /109.47/ (0.195)
$p^{\#}\to\infty$	σ [NW s.e.] /F/ (J p-value)	0.464 [0.065] /13.77/	0.355 [0.095] /13.69/	0.414 [0.016]	0.460 [0.116] /92.09/	0.493 [0.107] /34.40/ (0.088)	0.459 [0.073] /57.37/ (0.141)

Table notes are placed after the final table in the manuscript.

Appendix VII:

Relating Heterogeneous Industry And Aggregated σ's

This appendix develops the formula for relating the aggregated σ (σ_{agg}) to heterogeneous industry σ's (σ_i). It then considers two alternative weighting schemes.

We begin with the definitions of σ_{agg} and σ_i that follow from equation (4) and are stated in terms of percentage changes in the aggregate and industry capital/output ratios ((K_{agg}/Y_{agg}) and (K_i/Y_i), respectively) and the aggregate and industry relative prices of capital (P_{agg}^{KY} and P_i^{KY}, respectively),

$$\sigma_{agg} \equiv \frac{d\left(K_{agg}/Y_{agg}\right)/\left(K_{agg}/Y_{agg}\right)}{d\,P_{agg}^{KY}/P_{agg}^{KY}}, \qquad Y_{agg} = \text{constant}, \qquad \text{(VII-1)}$$

$$\sigma_i \equiv \frac{d\left(K_i/Y_i\right)/\left(K_i/Y_i\right)}{d\,P_i^{KY}/P_i^{KY}}, \qquad Y_i = \text{constant}. \qquad \text{(VII-2)}$$

While the relative price of capital varies by industry, we consider a percentage change that is equal across all industries (e.g., a change in the nominal or relative price of investment),

$$d\,P_{agg}^{KY}/P_{agg}^{KY} = d\,P_i^{KY}/P_i^{KY} = d\,P/P, \qquad \forall i. \qquad \text{(VII-3)}$$

We begin with identities relating changes in the aggregated to industry capital, and changes in the aggregated and industry capital/output ratios, respectively,

$$\frac{dK_{agg}}{d\,P/P} \equiv \sum_i \frac{dK_i}{d\,P/P}, \qquad \text{(VII-4)}$$

$$\frac{dK_{agg}}{d\,P/P} \equiv \frac{d\left(K_{agg}/Y_{agg}\right)}{d\,P/P} * Y_{agg}, \qquad Y_{agg} = \text{constant}, \qquad \text{(VII-5)}$$

$$\frac{dK_i}{d\,P/P} \equiv \frac{d\left(K_i/Y_i\right)}{d\,P/P} * Y_i, \qquad Y_i = \text{constant}. \qquad \text{(VII-6)}$$

Substituting equations (VII-5) and (VII-6) into equation (VII-4), dividing both sides by K_{agg}, and rearranging, we obtain the following equation,

$$\frac{d\left(K_{agg}/Y_{agg}\right)/\left(K_{agg}/Y_{agg}\right)}{d\,P/P} = \sum_i \frac{d\left(K_i/Y_i\right)}{d\,P/P} * \left(Y_i/K_{agg}\right). \qquad \text{(VII-7)}$$

The left-side equals σ_{agg} by equation (VII-1). Multiplying the right-side by (K_i / K_i) and rearranging, we obtain the following equation,

$$\sigma_{agg} = \sum_i \frac{d(K_i / Y_i) * (K_i / Y_i)}{dP / P} * (K_i / K_{agg}).\qquad\qquad\text{(VII-8)}$$

Using the definition of σ_i from equation (VII-2) and defining the latter object in equation (VII-8) as an industry weight, we obtain the following equation,[6]

$$\sigma_{agg} = \sum_i \sigma_i * w_i, \qquad\qquad w_i \equiv (K_i / K_{agg}).\qquad\qquad\text{(VII-9)}$$

In equation (VII-9), the aggregated σ (σ_{agg}) is a weighted average of the industry σ's (σ_i's), where the w_i's are industry weights defined in terms of industry capital ratios.[7]

Alternative aggregation procedures have been developed in two papers. First, Oberfield and Raval (2014; hereafter, OR) present a scheme that estimates σ's at the plant level, and then aggregates them to the industry and manufacturing levels. Their procedure allows for reallocation of resources as a result of the change in factor prices.

The aggregation procedure developed in this Appendix and used in Section V differs from that of OR in that we evaluate our derivatives at the initial, pre-change point (see equations (VII-5) and (VII-6)), while OR evaluate derivatives at the post-change point. See the equation between equations (5) and (6) in their paper. Their $d\theta_i / d\ln(w / r)$ term represents the change in industry weights that is non-zero in OR and zero in our paper. (In a loose sense, the difference in aggregation procedures is akin to the difference between using a Paasche (OR) or Laspeyres (our paper) price indices.) The benefit of the OR procedure is that it allows for reallocations of factors and economic activity. However, their procedure also requires estimates of the own-industry and cross-industries demand elasticities and the elasticity of substitution between

---

[6] Note that σ can be defined in terms of the capital/labor ratio analyzed in Section IV.F. In this case, the derivation presented in this appendix is unaffected (merely replace Y_{agg} with L_{agg} and Y_i with L_i in equations (VII-5) to (VII-8) and P_{agg}^{KY} with P_{agg}^{KL} and P_i^{KY} with P_i^{KL} in equation (VII-3).

[7] It should be noted that the σ from the benchmark model is also effectively a weighted-average estimate. The heterogeneous model analyzed in this appendix weights the σ_i's by industry capital shares, while the homogeneous model effectively weights the σ_i's by relative industry variances.

material and non-material factors. Also, once changes in product demand are introduced, the price elasticity of demand for a factor of production is no longer equal to just σ. Rather, per Hicks formula for the derived demand of a factor of production (see Chirinko and Mallick, 2011, equations (10) or (11) for a recent statement of the Hicks formula), the price elasticity also depends on the capital income share and the own-demand elasticity.

A less important difference is that OR analyze σ's in terms of the capital/labor ratio, while we analyze the capital/output ratio.

Second, in a provocative paper, Jones (2005) formally relates industry and aggregate (global) production functions to the distribution of alternative production techniques (APT's) for combining capital and labor. His striking result is that the industry and aggregate production functions will be Cobb-Douglas in the long-run. This approach has the benefit of developing solid microfoundations for production functions but is sensitive to the assumed distribution of ideas. When APT's are distributed according to a Pareto distribution with independence between marginal APT distributions, the Cobb-Douglas result obtains. However, when APT's are distributed according to a Weibull distribution (Growiec (2008a)) or a Pareto distribution with dependence between marginal APT distributions (Growiec (2008b)), the industry and aggregate production functions are CES. These theoretical results, coupled with the empirical results presented in this paper, suggest the need for further study of aggregation procedures and the underlying distribution of ideas.

Appendix VIII:

Relating The Capital Share Of Income And σ_{agg}

This appendix formally relates the capital share of income to σ_{agg} in order to evaluate

the impact of a rising σ_{agg} in Figure 4 (from 0.63 to 0.67) on a rising capital share of income

(from 0.25 to 0.40; Chirinko, Wilson, Zidar, 2015, Slide 5). Time and firm subscripts, "*"

superscripts, and technology shocks are omitted for notational convenience.

The capital share of income (KS) is defined as follows,

$$KS \equiv \frac{MPK * K}{Y} \ . \qquad \text{(VIII-1)}$$

Equating the MPK to the relative price of capital (P^K / P^Y), denoting this relative price by the

the Jorgensonian user cost of capital (C), and substituting equation (2) for the K/Y ratio, we

obtain the following equation,

$$KS \ = \ C \, \phi^{\sigma} \, C^{-\sigma} \ = \ C^{1-\sigma} \, \phi^{\sigma} \ . \qquad \text{(VIII-2)}$$

Calibrating equation (VIII-2) by assuming that KS = 0.25 (the value in 1980 for the gross capital

share), C = 0.15, and $\sigma = 0.63$ (the value in 1980 from Figure 4), $\phi = 0.34$. If σ rises to 0.67,

KS rises by only one percentage point to 0.26.

References Cited Only In The Appendices

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

Acemoglu, Daron, *Introduction to Modern Economic Growth* (Princeton and Oxford: Princeton University Press, 2009).

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

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.

Chirinko, Robert S., and Mallick, Debdulal, "The Elasticity of Derived Demand, Factor Substitution, and Product Demand: Corrections to Hicks' Formula and Marshall's Four Rules" (with Debdulal Mallick), *Labour Economics* 18 (October 2011), 708-711.

Chirinko, Robert S., Wilson, Daniel J., and Zidar, Owen, "The Declining Labor Share: The View from the States," Presentation at the 2015 American Economic Association meetings.

Eicher, Theo S., and Turnovsky, Stephen J., "Non-Scale Models of Economic Growth," *Economic Journal* 109 (July 1999), 394-415.

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

Growiec, Jakub, "Production Functions and Distributions of Unit Factor Productivities: Uncovering the Link," *Economics Letters* 101 (2008a), 87-90.

Growiec, Jakub, "A New Class of Production Functions and an Argument Against Purely Labor-Augmenting Technical Change," *International Journal of Economic Theory* 4 (2008b), 483-502.

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

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

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

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

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.

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

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.

Turnovsky, Stephen J., and Smith, William T., "Equilbrium Consumption and Precautionary Savings in a Stochastically Growing Economy," *Journal of Economic Dynamics & Control* 30 (February 2006), 243-278.

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.