Is the Long-Run Labor Supply Vertical?

Paul Evans[a]

Georgios Karras[b,c]

June 2000

Abstract

Growth theory and real-business-cycle theory virtually always assume that the long-run labor supply is vertical. This paper derives the restriction on preferences required for this assumption to hold and formulates an econometric method for testing this restriction. The restriction is then rejected using time-series data for the United States and many OECD countries. Some evidence against the restriction is also found using cross-sectional data for 133 countries from the Penn World Tables. The long-run labor supply therefore appears not to be vertical for many countries.

Keywords: labor supply, growth theory, real business cycles, cointegration

***JEL* Codes:** O40, E32, J23, C22

---

[a]Address: Department of Economics, Ohio State University, Columbus, OH 43210; Email: evans21@osu.edu; Phone: 614-292-0072; Fax: 614-292-3906.

[b]Address: Department of Economics, University of Illinois at Chicago, Chicago, IL 60880; Email: gkarras@uic.edu; Phone: 312-996-2321; Fax: 312-996-3344.

[c]We thank the participants in seminars at Ohio State University, the University of Houston, Texas A&M, and Southern Methodist University for helpful comments. Especially helpful comments came from Mario Crucini and Nathan Balke.

I. Introduction

In the growth and real-business-cycle literatures, preferences are nearly always restricted so that the fraction of hours supplied as labor is independent of the wage rate when the wage rate grows at the same rate as per capita consumption. In other words, these literatures impose a vertical long-run labor supply along which the substitution and income effects from balanced changes in the wage rate exactly cancel each other.[1]

A vertical long-run labor supply is imposed for at least five reasons. First, this restriction seems natural for series bounded between zero and one. Second, it is widely believed that the fraction of hours supplied as labor has reverted toward a constant mean while the wage rate and per capita consumption have trended upward; see Prescott (1986, p.23) for an example. Third, a vertical long-run labor supply is required for balanced growth; see Barro and Sala-i-Martin (1995). Fourth, the conditions that define balanced growth have been widely accepted since Kaldor (1961) first stated them; see Romer (1996, p.15). Fifth, the nonexistence of balanced growth would typically wreak havoc on growth theory and would require substantial changes in the real-business-cycle literature. It is, then, natural in the absence of compelling evidence to the contrary to cling to assumptions that permit balanced growth.

It is a fallacy that series bounded between zero and one can neither trend nor be nonstationary. A simple counterexample is $x/(1+x)$, where x is any trended or difference-stationary series. A correct claim is that series bounded between zero and one cannot trend at a *constant* nonzero rate and do not have *linear* difference-stationary representations with *covariance-stationary* innovations. The logistic of such a series, however, can trend at a constant rate and can have such a representation. The formal empirical analysis of this paper employs the logistic transformation for this very reason.

[1] Indeed, among the many papers we have read in these literatures, we cannot remember ever seeing any that do not implicitly or explicitly impose a vertical long-run labor supply. A far from exhaustive list of recent articles that do is Aghion and Howitt (1992), Barro and Sala-i-Martin (1992, 1997), Benhabib, Rogerson, and Wright (1991), Burnside, Eichenbaum and Rebelo (1993), Campbell (1994), Christiano and Eichenbaum (1992), Cogley and Nason (1995), Cooley and Ohanian (1997), Jones (1995), King, Plosser and Rebelo (1988a, 1988b), Kydland and Prescott (1982), Lucas (1988), Mankiw, Romer and Weil (1992), Ohanian (1997), Rebelo (1991), and Romer (1986, 1990). See also the articles collected in Barro (1989) and Cooley (1995) and the books by Barro and Sala-i-Martin (1995) and Grossman and Helpman (1993).

Stylized facts might be defined as propositions that one wishes were true but unfortunately are not. Figures 1 and 2 provide some basis for such a cynical definition.[2] According to Figure 1, the fraction of hours supplied as labor in the United States fell sharply until the end of World War II, drifted gradually lower over the next thirty years, and then rose somewhat after the mid-1970s. The fall in this measure of labor supply has largely reflected shorter workdays and workweeks, more holidays and other time off work, longer vacations, later entry into the labor force, earlier retirement, and more years of retirement.[3] Figure 2 shows that the employment rate fell in most OECD countries over the period 1960-1994 but rose in a few. These series also appear to be difference stationary.

Figures 1 and 2 understate the evidence against a vertical long-run labor supply. First, the hours supplied to the labor market have no doubt grown increasingly less physically demanding and mentally tedious while the hours spent in the home have been increasingly devoted to leisure rather than work.[4] Second, female participation in the labor market has increased rapidly over the past century and especially during the postwar period. For the United States and most OECD countries, plots disaggregated by gender would show that the fraction of hours supplied has greatly increased for females and has appreciably declined for males; see Killingsworth and Heckman (1986) and Pencavel (1986).[5] Following the growth and real-business-cycle literatures, we abstract from such shifts, treating each household as if it were a unitary decision maker and assuming that if the long-run labor supply curve is indeed vertical, households have merely been substituting female for male labor.

---

[2] The series plotted in these figures are defined and described in sections III and IV.

[3] According to Table 3.1 of Cox and Alm (1999), the average workday declined from 9.7 hours in 1890 to 7.3 hours in 1990; the average workweek declined from 6.0 days in 1890 to 4.7 days in 1990; holidays increased from 3 in 1890 to 11 in 1990; average vacation days increased from 0 in 1890 to 10 in 1990; and other time off increased on average from 8 days in 1890 to 9.5 days in 1990. According to their Table 3.2, the average age of entry into the labor force increased from 13 years in 1870 to 17.6 years in 1950 and 19.1 years in 1990; workers typically did not retire before death in 1870 but retired on average at age 68.5 years in 1950 and 63.6 years in 1990; and retirement lasted an average of 3.5 years in 1950 and 12.9 years in 1990.

[4] Table 3.2 of Cox and Alm (1999) reports that annual hours spent working at home declined from 1825 in 1870 to 1544 in 1950 and 1290 in 1990.

[5]+ For example, if we had plotted the fraction of male hours supplied to the labor market, the downward trend exhibited in Figure 1 would have been appreciably more pronounced in the early postwar period and would not have been reversed after the mid-1970s.

The rest of the paper formally investigates whether the long-run labor supply is vertical. Section II derives the restriction on preferences required for it to be vertical and formulates an econometric method for testing whether this restriction holds. Sections III, IV, and V use data for the United States, 22 OECD countries, and 133 countries from the Penn World Tables to test whether the long-run labor supply is vertical, finding considerable evidence that it is not. Finally, section VI concludes.

II. Theoretical and Econometric Discussion

In this section, we lay out a simple theoretical model of labor supply. Our purpose is to derive the restriction on preferences necessary for the fraction of hours supplied as labor to be mean stationary in an environment in which per capita consumption and the wage rate have a common stochastic trend. In addition, we formulate a method for testing this restriction.

A. Theory

We assume that the economy consists of a unit mass of identical infinitely-lived households, each of which consists of a constant number of individuals.[6] Each of these households maximizes an objective function of the form

$$E_0 \sum_{t=0}^{\infty} (1+\rho)^{-t}\, \mathrm{U}(C_t, L_t, z_t), \quad \rho > 0, \tag{1}$$

subject to a sequence of budget constraints of the form

$$C_t + A_{t+1} \leq W_t L_t + R_t A_t, \quad t = 1, 2, 3, \cdots, \tag{2}$$

a sequence of constraints on indebtedness, and an initial condition on A_0. In (1) and (2), t is a discrete index of time; C_t is the per capita consumption of the household during period t; L_t is the fraction of the available time that members of the household supply to

[6] We are inclined to agree with Killingsworth and Heckman (1986) and Pencavel (1986), who argue that the assumption of *ex ante* identical households (or of a representative household) is indefensible and cannot be expected to provide useful predictions about the aggregate labor supply. We rationalize this assumption on the grounds that it is ubiquitous in the literatures that we seek to criticize.

the labor market; z_t is a vector of exogenous preference shifters;[7] A_t is the stock of assets held by the household at the beginning of period t; W_t and R_t are the wage and gross rental rates during period t; U(•) is a twice continuously differentiable momentary utility function that is increasing in its first argument, decreasing in its second argument, and strictly concave in both arguments; and ρ is the subjective discount rate.

One of the first-order conditions for this problem is[8]

$$-\mathrm{U}_L(C_t, L_t, z_t) = W_t\, \mathrm{U}_C(C_t, L_t, z_t) \tag{3}$$

where the subscripts on U(•) denote derivatives. Condition (3) merely states that each household equates its marginal disutility of work to its marginal utility from consuming the wage. This condition holds whether the constraints on indebtedness merely require the household to be solvent or hinder it from borrowing against future wage income. By contrast, the condition would not hold, were either consumption or labor supply durable or habit-forming. For our purposes, however, we think that treating them as if they are completely perishable is a reasonable approximation.[9]

Taking logarithms of both members of equation (3) and rearranging yields

$$\ln(W_t L_t / C_t) = \ln[-L_t\, \mathrm{U}_L(C_t, L_t, z_t)] - \ln[C_t\, \mathrm{U}_C(C_t, L_t, z_t)]. \tag{4}$$

It is convenient to define the function

$$u(c, l, z) \equiv U[\mathrm{e}^c, \mathrm{e}^l/(1+\mathrm{e}^l), z],$$

where $c \equiv \ln C$ and $l \equiv \ln[L/(1-L)]$. Note that l, the logistic of L, lies on $\Re$ since U(•) is assumed to have properties that lead L to lie on the interval (0,1). With this definition, we can rewrite equation (4) in the form

$$\ln(W_t L_t / C_t) = \ln[-\mathrm{u}_l(c_t, l_t, z_t)] - \ln[1+\exp(l_t)] - \ln[\mathrm{u}_c(c_t, l_t, z_t)]. \tag{5}$$

---

[7] We could also include choice variables in z. Doing so would complicate the analysis without producing any added insights.

[8] We are assuming here that an interior solution exists; that is, a solution for which $C > 0$ and $0 < L < 1$. Imposing the restriction $L > 0$ is natural since we are assuming that all households are identical *ex ante*. In fact, some households choose never to work, a key criticism that labor economists make of representative-agent models. See Killingsworth and Heckman (1986) and Pencavel (1986) for further discussion.

[9] For example, suppose that $\ln C_t$ is difference stationary and that momentary utility depends on the quantity $C_t+\theta_1 C_{t-1}+\theta_2 C_{t-2}+\cdots$, where the θs are parameters. Then the logarithm of this quantity is cointegrated with $\ln C_t$ itself.

Finally, linearizing equation (5) around the reference values $c_t = \breve{c}$, $l_t = \breve{l}$, and $z_t = 0$, we obtain

$$\ln(W_t L_t / C_t) = \alpha + \beta c_t + \gamma l_t + \zeta' z_t, \qquad (6)$$

where

$$\beta \equiv u_{cl}(\breve{c},\breve{l},0)/u_l(\breve{c},\breve{l},0) - u_{cc}(\breve{c},\breve{l},0)/u_c(\breve{c},\breve{l},0),$$

$$\gamma \equiv u_{ll}(\breve{c},\breve{l},0)/u_l(\breve{c},\breve{l},0) - u_{cl}(\breve{c},\breve{l},0)/u_c(\breve{c},\breve{l},0) - \exp(\breve{l})/[1+\exp(\breve{l})],$$

$$\zeta \equiv u_{lz}(\breve{c},\breve{l},0)/u_l(\breve{c},\breve{l},0) - u_{cz}(\breve{c},\breve{l},0)/u_c(\breve{c},\breve{l},0),$$

and

$$\alpha \equiv \ln[-u_l(\breve{c},\breve{l},0)] - \ln[1+\exp(\breve{l})] - \ln[u_c(\breve{c},\breve{l},0)] - \beta\breve{c} - \gamma\breve{l}.$$

Suppose that the logarithm of per capita consumption is difference stationary with a nonzero drift rate. A balanced stochastic growth path then exists only if the logarithm of the ratio of wage income to consumption is mean stationary and the logistic of the fraction of time supplied as labor is mean stationary. Equation (6) then implies that except by happenstance, a balanced stochastic growth path can exist only if (a) $\zeta' z_t$ is mean stationary and (b) $\beta = 0$.[10] We label condition (b) as the null hypothesis H_0: $\beta = 0$, which we seek to test against the alternative hypothesis H_1: $\beta \neq 0$. The next subsection shows that such a test can be performed provided that condition (a) can be maintained as an assumption. It must be maintained because nonstationary movement along a nonvertical long-run labor supply cannot be readily distinguished from nonstationary movement along a vertical long-run labor supply that is undergoing nonstationary shifts in location.

Households may conceivably be off their static labor supply relationships. In particular, the observed quantity of labor exchanged during recessions, wars, and especially the Great Depression may not satisfy equation (6), whose derivation assumed

---

[10] It is straightforward to show that $\beta = 0$ implies $CU_{CC}/U_C - CU_{CL}/U_L = 1$. In order for the mean growth rate of per capita consumption to be constant, $-CU_{CC}/U_C$ must also be constant. Barro and Sala-i-Martin (1995, pp. 326-327) have shown that these two conditions can hold if, and only if, the momentary utility function takes the form

$$H(z) + \{[C\,V(L,z)]^{1-\omega} - 1\}/(1-\omega), \quad \omega > 0,$$

where ω is a parameter, $V(\bullet)$ is a strictly decreasing and strictly concave function, and $H(\bullet)$ is any function.

no search costs or other frictions in the labor market.[11] We allow for this possibility by assuming that the observed quantity of labor exchanged is the quantity implied by the theoretical model plus a mean-stationary error term. We therefore have

$$\ln(W_t L_t / C_t) = \alpha + \beta c_t + \gamma l_t + e_t, \qquad (7)$$

where e_t is $\zeta' z_t$ plus 1-γ times this mean-stationary error term.

B. Econometrics

For the moment, suppose that e_t is covariance stationary with a zero mean. Under the alternative hypothesis that $\beta \neq 0$, equation (6) and the difference stationarity of c_t imply that $\ln(W_t L_t / C_t)$ or l_t or both are difference stationary. Except by happenstance, this departure from balanced growth manifests itself in both $\ln(W_t L_t / C_t)$ and l_t, and these two variables are not themselves cointegrated. As a result, $\ln(W_t L_t / C_t)$, c_t and l_t are cointegrated with the cointegrating vector (1,-*β*,-*γ*). The ordinary least squares estimator of β is therefore superconsistent. Furthermore, inference based on the *t*-ratio of this estimator is entirely standard if the error term e_t is serially uncorrelated, homoskedastic, and uncorrelated with all leads and lags of Δc_t and Δl_t; for proof, see Hamilton (1994, pp. 602-608). Unfortunately, these restrictions are unlikely to be satisfied in practice.

Correlation of e_t with Δc_t, Δl_t, or their leads and lags can be handled by specifying e_t to take the form

$$e_t = \sum_{j=m}^{-m} \lambda_j \Delta c_{t-j} + \sum_{j=m}^{-m} \mu_j \Delta l_{t-j} + \varepsilon_t, \qquad (8)$$

where the λs and μs are parameters and *m* is an integer sufficiently large to eliminate essentially all correlation between the error term ε_t and Δc_t, Δl_t, or their leads and lags. Substituting equation (7) into equation (8) gives us

[11] For example, in Figure 1, *L* for the United States fell sharply during the contractions of the 1890s, the early 1920s, and the 1930s and rose appreciably during World War II.

$$\ln(W_t L_t / C_t) = \alpha + \beta c_t + \gamma l_t + \sum_{j=m}^{-m} \lambda_j \Delta c_{t-j} + \sum_{j=m}^{-m} \mu_j \Delta l_{t-j} + \varepsilon_t. \tag{9}$$

Ordinary least squares yields consistent estimators of the parameters of equation (9). Consequently, on the assumption that ε_t follows the pth-order autoregression

$$\varepsilon_t = \sum_{k=1}^{p} \phi_k \varepsilon_{t-k} + u_t, \tag{10}$$

the parameters $\phi_1, \phi_2, \cdots, \phi_p$ can be estimated consistently by regressing the ordinary-least-squares residuals from equation (9) on p lags of themselves. Let $\hat{\phi}(L) \equiv 1 - \sum_{k=1}^{p} \hat{\phi}_k L^k$, where L is the lag operator and $\hat{\phi}_1, \hat{\phi}_2, \cdots, \hat{\phi}_p$ are the estimators so obtained. It then follows immediately that the error term υ_t in the regression equation

$$\begin{aligned} \hat{\phi}(L)\ln(W_t L_t / C_t) &= \alpha + \beta[\hat{\phi}(L)c_t] + \gamma[\hat{\phi}(L)l_t] \\ &+ \sum_{j=m}^{-m} \lambda_j[\hat{\phi}(L)\Delta c_{t-j}] + \sum_{j=m}^{-m} \mu_j[\hat{\phi}(L)\Delta l_{t-j}] + \upsilon_t \end{aligned} \tag{11}$$

is asymptotically equivalent to the error term u_t in equation (10). Therefore, applying ordinary least squares with White's (1980) correction for heteroskedasticity should produce not only a superconsistent estimator of β but also consistent estimators of the t-ratio and standard error. For further discussion of the method described here, see Hamilton (1994, pp. 608-612).

Allowing for the possibility that e_t is difference stationary rather than mean stationary is important for two reasons. First, preferences may change permanently, leading e_t to be difference stationary. For example, the postwar entry of females into the labor market might represent such a preference shift. This possibility arises whether $\beta = 0$ or $\beta \neq 0$. Second, if $\beta \neq 0$, the linearization of equation (5) that we employed in obtaining (6) omits a remainder term that is nonstationary and therefore cannot be neglected since it is quadratic in the difference-stationary variable c_t.[12]

If e_t is difference stationary rather than mean stationary, the estimator described above converges in probability to zero only by happenstance, and its t-ratio diverges. As a

[12] Note that the remainder term does not depend on c_t if the momentary utility function takes the form in footnote 10. For this reason, it can be taken to be mean-stationary under the null hypothesis.

result, a sufficiently large sample should enable the null hypothesis to be rejected if either the long-run labor supply is not vertical or it shifts permanently over time. By contrast, failure to reject the null hypothesis indicates that the long-run labor supply is vertical and shifts only temporarily. Strictly speaking, then, the method described above cannot distinguish between a long-run labor supply that is not vertical and one that shifts permanently.[13] This distinction is not important in practice, however, since no balanced growth path exists in either case.

III. Empirical Results for the United States

In this section, we investigate whether the long-run labor supply is vertical. Our initial step is to examine whether plausible measures of the fraction of hours supplied as labor are better characterized as mean stationary rather than either trending or difference stationary.

We consider the following two annual measures of the labor supply L: (i) civilian employment times annual hours worked per employee in the private sector divided by the mid-year civilian population between the ages of 16 and 64 times the maximum number of hours possible per year; and (ii) a variant of (i) in which the civilian labor force replaces civilian employment.[14] Our estimate of the maximum number of hours per year is 16 hours per day times 7 days per week times 52 weeks per year, or 5824 hours per year. Measure (i) is appropriate if time spent searching for employment is equivalent to leisure, while measure (ii) is appropriate if search time is equivalent to work.[15] We consider both measures because the truth is likely to lie somewhere between these two extremes.

---

[13] It is possible in principle to distinguish the two when e_t is difference stationary. The parameter β can be estimated consistently by differencing equation (6) and fitting the resulting equation using valid instrumental variables. Unfortunately, identifying instrumental variables that are well correlated with Δc_t and Δl_t but uncorrelated with Δe_t is a daunting task. We avoid this task using the method described in the text, but only with the cost of not being able to distinguish a long-run labor supply that is not vertical from one that shifts permanently over time.

[14] We also considered measures that include the military in employment and the labor force, obtaining similar results.

In addition to these two measures of the per capita labor supply, we examine the following measures for C and WL/C: (iii) the annual real consumption of nondurable goods and services divided by the population aged 16-64; and (iv) after-tax annual nominal wage income divided by annual nominal consumption of nondurable goods and services. For further description of these four series and their sources, see the appendix.

We first apply the augmented Dickey-Fuller unit-root test to the logistics of series (i) and (ii) and the logarithms of series (iii) and (iv). Table 1 reports statistics for the longest feasible sample period and for the postwar sample period 1947-1995. We report only the results for an augmentation lag of two years; the results for lags ranging from zero to five years are similar. The reported statistics provide little evidence against the null hypothesis of difference stationarity. On the hypothesis that these series are indeed difference stationary, Table 2 provides little evidence that either labor-supply measure trends over either sample period. By contrast, the evidence that per capita consumption trends is compelling. Finally, some evidence exists for a trend in the ratio of wage income to consumption over the period 1947-1995, but none exists for the longer period 1930-1995.[16]

The univariate time-series properties of our two labor supply measures provide no evidence for a vertical long-run labor supply. One should not, however, regard this evidence as more than suggestive since a more powerful test is available. On the maintained hypothesis that the logarithm of per capita consumption is difference stationary, we can directly test whether the long-run labor supply is vertical using the method formulated in section II. Following this method produced the estimates and standard errors reported in Table 3. All but two of the estimates are statistically significant at the .001 level, and those two are significant at the .05 and .10 levels. We

[15] Measured unemployment equals the gap between labor supplied and exchanged only if the latter assumption holds.

[16] Canjels and Watson (1997) establish that the *t*-ratios of the estimates reported in Table 2 need not converge in distribution to standard normal if the series are trend rather than difference stationary. For that reason, we also investigated whether each of their trends is significant conditional on the hypothesis of trend stationarity with a dominant autoregressive root of ψ. Using a Newey-West correction for heteroskedasticity and additional serial correlation with a window length of five years, we obtained results similar to those reported in the text except for series (ii) over the sample period 1900-1995. In this case, significant evidence was found for a downward trend if $\psi \leq .92$ but no evidence if $.93 \leq \psi < 1$.

can therefore strongly reject the null hypothesis in favor of the alternative hypothesis. Apparently, then, the long-run labor supply for the United States is not vertical.

IV. Empirical Results for a Sample of OECD Countries

We now turn to a sample of 22 OECD countries for which data on labor force participation and employment are available for the entire period 1960-1994.

First, we examine the univariate properties of the labor supply using the following four measures: (i) total labor force divided by the population aged 15-64; (ii) a variant of (i) in which total employment replaces total labor force; (iii) total labor force multiplied by an index of hours per employee worked in manufacturing divided by the population aged 15-64; and (iv) a variant of (iii) in which total employment replaces total labor force.[17] Series (i) and (ii) are available for the entire sample, but series (iii) and (iv) can only be constructed for the seven countries for which hours data are available for each year during 1962-1993: Canada, Germany, Greece, Ireland, Japan, the United Kingdom, and the United States.

In addition to these three series, we examine the following measures for *C* and *WL/C*: (v) annual real consumption divided by the population between the ages of 15 and 64;[18] and (vi) annual total pretax nominal employee compensation divided by nominal consumption.[19] Data sources and definitions can be found in the appendix.

Table 4 reports augmented Dickey-Fuller statistics for the logistics of series (i)-(iv) and the logarithms of series (v) and (vi) for the 22 countries. We set the

[17] Note that because only hours indices are available for some of the countries, we multiply by an index of hours rather than the ratio of hours to maximum hours. As a result, the labor supply measures used here are not exact analogues of those used in section II.

[18] We should prefer to use consumption of nondurable goods and services as in section III, but the data are available for only a few countries. As pointed out in n. 6, however, using a measure of consumption that is somewhat durable may not seriously misspecify the model.

[19] We use a pretax rather than an after-tax measure of wage income because an after-tax measure can be constructed for only ten countries and because we do not know the OECD tax systems well enough to be confident that we have constructed reasonably valid measures. We did, however, perform the statistical analyses below using an after-tax measure for these ten countries, obtaining similar results. Note that along a balanced growth path, the ratio of wage taxes to consumption is mean stationary so that the use of a pretax rather than an after-tax measure contributes only mean-stationary measurement errors to ln(*WL/C*). For this reason, the method described in section II for making inferences about β remains valid.

augmentation lag length equal to one; results for other lag lengths are similar. The evidence for most countries is convincingly consistent with the null hypothesis of difference stationarity for all six series. On the hypothesis that the series are indeed difference stationary, Table 5 shows that for the overwhelming majority of countries, there is no evidence of a trend in any of the four labor-supply measures or in the ratio of wage income to consumption. Per capita consumption, on the contrary, has statistically significant trends for all but three of the countries.

Tables 6 and 7 conduct unit-root tests for the series (i)-(vi), considering the data as a panel. Table 6 uses the procedure outlined in Levin and Lin (1992). For each of the series, the null hypothesis of difference stationarity cannot be rejected for any of the lag lengths tried. Similar results emerge when we use the Im-Pesaran-Shin unit root test in Table 7. This statistic is calculated as the cross-country average of $(DF_i - \mu_{DF})/\sigma_{DF}$, where DF_i is the Dickey-Fuller *t*-ratio for country *i* and μ_{DF} and σ_{DF} are the mean and standard deviation of the *t*-ratio from the Dickey-Fuller distribution under the null hypothesis of a unit root.[20] Consistent with the Levin-Lin procedure, the Im-Pesaran- Shin technique cannot typically reject the null of nonstationarity for series (i)-(v). The results of the two methods differ, however, for series (vi), the logarithm of the ratio of wage income to consumption, which Table 7 indicates to be mean stationary.

Table 8 reports estimates of β for each of the labor supply measures (i)-(iv) and for the 22 OECD countries. Pretesting indicated that $p = 3$ is adequate to whiten the error terms. We included $m = 0, 1, 2$, and 3 leads and lags of Δc and Δl in the regressions, obtaining similar results in all cases. Only the results for $m = 1$ are reported. Most of the estimates are highly statistically significant, offering strong evidence against the null hypothesis and implying that most of the countries lack vertical long-run labor supplies.

V. Some Cross-Sectional Results for Larger Samples

[20] A Monte Carlo simulation with 100,000 iterations gave $\mu_{DF} = -1.52115$ and $\sigma_{DF} = 0.89009$ for models without trend, which we used for series (i)-(iv) and (vi), and $\mu_{DF} = -2.17723$ and $\sigma_{DF} = 0.82643$ for models with trend, which we used for series (v).

This section utilizes information for samples from the Penn World Tables 5.6 (documented in Summers and Heston, 1991) in order to estimate a number of cross-sectional versions of our empirical model. We do not estimate equation (11) directly for this sample because we have no data on wage income for most of the PWT countries and we are reluctant to rely on annual observations, many of which are interpolated. Instead, we proxy wage income with gross domestic product and rely only on observations from the years 1950, 1960, 1970, 1980, and 1990, several of which are benchmark years. On the assumption that the parameters β and γ do not vary across countries, we can rewrite equation (7) in the form

$$l_{it} = -\alpha_i / \gamma + \delta_1 \ln(W_{it} L_{it} / C_{it}) + \delta_2 c_{it} + v_{it},$$

where i indexes over countries, $\delta_1 \equiv 1/\gamma$, $\delta_2 \equiv -\beta/\gamma$, and $\beta_t \equiv$ [illegible]. Here, we allow each country to have its own fixed effect α_i. Differencing the fixed effects out, we then obtain

$$l_{iT} - l_{iS} = \delta_0 + \delta_1[\ln(Y_{iT} / C_{iT}) - \ln(Y_{iS} / C_{iS})] + \delta_2 (c_{iT} - c_{iS}) + w_{it}, \qquad (12)$$

where T and S are two distinct years, l is the logistic of employment participation per equivalent adult, Y is real GDP, C is real consumption, c is the logarithm of consump-tion per equivalent adult, δ_0 is the cross-sectional mean of $\ln(W_{iT}L_{iT}/Y_{iT})$- $\ln(W_{iS}L_{iS}/Y_{iS})$, and $\mathrm{H}(z) + \{[C\,\mathrm{V}(L, z)]^{1-\omega} - 1\}/(1-\omega), \quad \omega > 0,$. We assume that w_i is uncorre-lated cross-sectionally as well as with the regressors. The latter is a strong assumption, the best justification of which may be necessity, as the task of finding appropriate instruments is even more formidable. Note that $\delta_2 = 0$ under the null hypothesis H_0: $\beta = 0$ and that $\delta_2 \neq 0$ under the alternative H_1: $\beta \neq 0$.

Table 9 reports the estimated δ_2s for ten pairs of years and for two sets of countries. The first set comprises all countries with data available over the relevant period, and the second set comprises the countries that have participated in the benchmark studies of the International Comparison Program (ICP) of the United Nations. All estimated δ_2s are positive, but the evidence on their statistical significance is mixed. In particular, the estimated parameters are statistically significant only for spans whose initial years are 1960 and 1970.

VI. Summary

One of the most frequently maintained assumptions in growth and real-business-cycle theories is that the long-run labor supply is vertical. Using a theoretical model for the determination of the optimal labor supply, the paper derives the restriction on preferences required for this assumption to be valid. If this restriction holds and if the other conditions required for the existence of a balanced growth path also hold, the logistic l of the fraction of time supplied as labor is mean stationary. By contrast, if the restriction does not hold and if the logarithm c of consumption is difference stationary, the logarithm of the ratio of wage income to consumption is cointegrated with c and l with a nonzero coefficient on c. Whether the long-run labor supply is vertical can therefore be tested by examining whether l is mean stationary or by regressing the logarithm of the ratio of wage income to consumption on c and l and testing whether the coefficient on c is indeed zero. Performing these tests on data for the U.S., 22 OECD countries, and 133 countries from the Penn World Tables, the paper finds strong evidence that the long-run labor supply is not vertical for many of these countries.

The most obvious implication of this finding is that the ubiquitous assumption of a vertical long-run labor supply with a mean-stationary location—despite the modeling conveniences that it affords—may be much less innocuous than is generally thought. In particular, our finding points to the possibility that over time work effort will continue to decline (perhaps asymptotically to zero), suggesting that growth theorists should emulate those science fiction writers who have imagined a future in which people do not work much if at all. For real-business-cycle theorists, a non-vertical labor supply means that specifying the intertemporal elasticity of substitution for consumption and the fraction of time supplied as labor do not suffice for calibration; a labor-supply elasticity needs to be specified, too.

For applied empirical work, an important implication of our findings is related to the process of convergence in per capita income across countries, a topic of considerable recent research interest. If the long-run labor supply is not vertical and work effort declines as countries become richer, the actual growth rate will be lower on average than its steady-state value because of the gradual withdrawal of labor from the market, affecting the relationship between starting income and growth and thus the observed speed of convergence. Moreover, our findings suggest that studies using per capita income are not as reliable as those using income per worker, as the former implicitly assume a vertical long-run labor supply. Finally, those empirical studies that have implicitly or explicitly assumed the mean-stationarity of the labor supply have produced inconsistent, and perhaps seriously biased, estimates.

References

Aghion, Philippe, and Howit, Peter. "A Model of Growth through Creative Destruction." *Econometrica* 60 (March 1992): 323-351.

Barro, Robert J. *Modern Business Cycle Theory*. Cambridge, MA: Harvard University Press, 1989.

Barro, Robert J., and Sala-i-Martin, Xavier. "Convergence." *Journal of Political Economy* 100 (April 1992): 223-251.

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

Barro, Robert J., and Sala-i-Martin, Xavier. "Technological Diffusion, Convergence, and Growth." *Journal of Economic Growth* 2 (March 1997): 1-26.

Benhabib, Jess; Rogerson, Richard; and Wright, Randall. "Homework in Macroeconomics: Household Production and Aggregate Fluctuations." *Journal of Political Economy* 99 (December 1991): 1166-1187.

Burnside, Craig; Eichenbaum, Martin; and Rebelo, Sergio. "Labor Hoarding and the Business Cycle." *Journal of Political Economy* 101 (April 1993): 245-273.

Campbell, John Y. "Inspecting the Mechanism: An Analytical Approach to the Stochastic Growth Model." *Journal of Monetary Economics* 33 (June 1994): 463-508.

Christiano, Lawrence J., and Eichenbaum, Martin. "Current Real-Business-Cycle Theories and Aggregate Labor-Market Fluctuations." *American Economic Review* 82 (June 1992): 430-450.

Cogley, Timothy, and Nason, James M. "Output Dynamics in Real-Business-Cycle Models." *American Economic Review* 85 (June 1995): 492-511.

Cooley, Thomas F., ed. *Frontiers of Business Cycle Research*. Princeton, NJ: Princeton University Press, 1995.

Cooley, Thomas F., and Ohanian, Lee E. "Postwar British Economic Growth and the Legacy of Keynes." *Journal of Political Economy* 105 (June 1997): 439-472.

Cox, W. Michael, and Alm, Richard. *Myths of Rich & Poor*. New York: Basic Books, 1999.

Grossman, Gene M., and Helpman, Elhanan. *Innovation and Growth in the Global Economy*. Cambridge, MA: MIT Press, 1993.

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

Im, Hyung So; Pesaran, M. Hashem; and Shin, Yongcheol. “Testing for Unit Roots in Heterogeneous Panels.” Mimeo, University of Cambridge, June 1995.

Jones, Charles I. “R&D-Based Models of Economic Growth.” *Journal of Political Economy* 103 (August 1995): 759-784.

Kaldor, Nicholas. “Capital Accumulation and Economic Growth.” In Lutz, F. A., and Hague, D. C., eds., *The Theory of Capital*. New York: St. Martin’s Press, 1961.

Killingsworth, Mark R., and Heckman, James J. “Female Labor Supply: A Survey.” In Ashenfelter, Orley C., and Layard, Richard, ed., *Handbook of Labor Economics, Volume 1*. Amsterdam: North-Holland, 1986.

King, Robert G.; Plosser, Charles I.; and Rebelo, Sergio. “Production, Growth and Business Cycles: I. The Basic Neoclassical Model.” *Journal of Monetary Economics* 21 (March/May 1988a): 195-232.

King, Robert G.; Plosser, Charles I.; and Rebelo, Sergio. “Production, Growth and Business Cycles: II. New Directions.” *Journal of Monetary Economics* 21 (March/May 1988b): 309-342.

Kydland, Finn E., and Prescott, Edward C. “Time to Build and Aggregate Fluctuations.” *Econometrica* 50 (November 1982): 1345-1370.

Levin, Andrew, and Lin, Chien-Fu. “Unit Root Tests in Panel Data: Asymptotic and Finite-Sample Properties.” Mimeo, University of California, San Diego, May 1992.

Lucas, Robert E., Jr. “On the Mechanics of Economic Development.” *Journal of Monetary Economics* 22 (July 1988): 3-42.

Mankiw, N. Gregory; Romer, David; and Weil, David N. “A Contribution to the Empirics of Economic Growth.” *Quarterly Journal of Economics* 107 (May 1992): 407-438.

Newey, Whitney K., and West, Kenneth D. “A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix.” *Econometrica* 55 (May 1987): 703-708.

Ohanian, Lee E. “The Macroeconomic Effects of War Finance in the United States: World War II and the Korean War.” *American Economic Review* 87 (March 1997): 23-40.

Pencavel, John. “Labor Supply of Men: A Survey.” In Ashenfelter, Orley C., and Layard, Richard, ed., *Handbook of Labor Economics, Volume 1*. Amsterdam: North-Holland, 1986.

Prescott, Edward C. “Theory Ahead of Business-Cycle Measurement.” *Carnegie-Rochester Conference Series on Public Policy* 25 (Autumn 1986): 11-44.

Rebelo, Sergio. “Long-Run Policy Analysis and Long-Run Growth.” *Journal of Political Economy* 99 (June 1991): 500-521.

Romer, David. *Advanced Macroeconomics*. New York: McGraw-Hill, 1996.

Romer, Paul M. “Increasing Returns and Long-Run Growth.” *Journal of Political Economy* 94 (October 1986): 1002-1037.

Romer, Paul M. “Endogenous Technological Change.” *Journal of Political Economy* 98 (October 1990): S71-S102.

Summers, Robert, and Heston, Alan. “The Penn World Tables (Mark 5): An Expanded Set of International Comparisons, 1950-1988.” *Quarterly Journal of Economics* 106 (May 1991): 327-368.

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

Table 1. Augmented Dickey-Fuller Statistics

Series	***Long Sample Period***	***Postwar Sample Period***
(i)	-2.24	-1.98
(ii)	-1.92	-0.72
(iii)	-2.24	-0.62
(iv)	-2.01	0.13

Notes. The statistics for series (i), (ii), and (iv) are for the null hypothesis of trendless difference stationarity against the alternative hypothesis of mean stationarity; those for series (iii) are for the null hypothesis of trending difference stationarity against the alternative hypothesis of trend stationarity. The long sample period is 1903-1995 for series (i) and (ii) and 1932-1995 for series (iii) and (iv); the postwar sample period is 1947-1995 for all series.

Table 2. Means and their Standard Errors for Changes in Series (i)-(iv)

Series	***Long Sample Period***	***Postwar Sample Period***
(i)	-.00278 (.00494)	.00055 (.00233)
(ii)	-.00274 (.00199)	.00103 (.00127)
(iii)	.0154[a] (.0045)	.0174[a] (.0025)
(iv)	-.00162 (.00516)	-.00469[b] (.00221)

Notes. The long sample period is 1901-1995 for series (i) and (ii) and 1930-1995 for series (iii) and (iv); the postwar period is 1947-1995 for all series. Standard errors appear in parentheses below their means and are calculated using a Newey-West correction for serial correlation and heteroskedasticity with a window length of five years. The superscripts a and b indicate statistical significance at the .001 and .05 levels, respectively.

Table 3. Estimates of β and their Standard Errors

for Two Labor-Supply Measures and Two Sample Periods

Lag and Lead Length (m)	*(i), Long Sample Periods*	*(ii), Long Sample Periods*	*(i), Postwar Sample Periods*	*(ii), Postwar Sample Periods*
0	-.266[a] (.039)	-.119[c] (.070)	-.314[a] (.037)	-.285[a] (.024)
1	-.294[a] (.070)	-.111[b] (.052)	-.331[a] (.027)	-.293[a] (.022)
2	-.214[a] (.062)	-.145[a] (.036)	-.323[a] (.031)	-.284[a] (.020)
3	-.206[a] (.044)	-.155[a] (.024)	-.308[a] (.034)	-.278[a] (.014)
4	-.233[a] (.038)	-.166[a] (.027)	-.273[a] (.036)	-.276[a] (.017)
5	-.268[a] (.035)	-.186[a] (.024)	-.204[a] (.039)	-.286[a] (.015)

Notes. The long sample periods extend from 1931+*m* to 1995-*m*; the short sample periods extend from 1947 to 1995-*m*. Standard errors appear in parentheses below their estimates. The superscripts a, b, and c indicate statistical significance at the .001 level, .05, and .10 levels, respectively.

Table 4. Augmented Dickey-Fuller Statistics

Country	***(i)***	***(ii)***	***(iii)***	***(iv)***	***(v)***	***(vi)***
Australia	-0.45	-3.05[b]			-2.25	-1.53
Austria	-2.38	-2.51			-1.13	-1.82
Belgium	-1.50	-1.12			-1.46	-3.04[b]
Canada	-0.52	-1.69	-2.36	-2.19	-2.24	-2.05
Denmark	-0.47	-3.30[b]			-2.20	-2.23
Finland	-1.58	-1.64			-1.06	-2.32
France	-2.57	-0.62			-1.66	-1.78
Germany	-2.51	-1.79	-0.49	0.46	-0.70	-2.18
Greece	-3.88[a]	-4.50[a]	-0.62	-0.58	-2.08	-2.10
Ireland	-2.41	-1.08	-1.07	-0.94	-1.76	-1.80
Italy	-3.54[b]	-3.03[b]			-1.75	-1.99
Japan	-2.24	-2.98	-1.64	-1.63	-2.48	-2.17
Netherlands	-1.45	-1.63			-2.68	-1.97
New Zealand	-2.17	-0.50			-1.63	-1.42
Norway	-1.14	-1.67			-1.42	-2.69
Portugal	-1.43	-1.47			-1.82	-2.18
Spain	-1.46	-0.56			-1.51	-2.52
Sweden	-1.62	-1.94			-1.56	-2.49
Switzerland	-2.82	-2.58			-3.11	-1.23
Turkey	-1.36	-2.25			-3.30	-0.24
United Kingdom	-1.82	-2.82	-1.14	-1.20	-2.38	-1.39
United States	0.35	-0.62	-0.10	-0.20	-2.23	-1.36

Notes. The sample period is 1964-1993. The statistics for series (i)-(iv) and (vi) are for the null hypothesis of trendless difference stationarity against the alternative of mean stationarity; those for series (v) are for the null hypothesis of trending difference stationarity against the null hypothesis of trend stationarity. The superscripts a and b indicate statistical significance at the .01 and .05 levels, respectively.

Table 5. Means of the Differences of Series (i)-(vi)

Country	*(i)*	*(ii)*	*(iii)*	*(iv)*	*(v)*	*(vi)*
Australia	.008	-.001			.018[a]	.000
Austria	-.003	-.005			.029[a]	.008
Belgium	.004	-.004			.025[a]	.008
Canada	.021[a]	.014[a]	.017[b]	.010	.020[a]	.005
Denmark	.015[b]	.000			.017[b]	.008
Finland	-.008	-.025			.024[a]	.008
France	-.006	-.015[a]			.025[a]	.005
Germany	-.001	-.008	-.021[a]	-.026[a]	.036[a]	.005
Greece	-.008	-.010	-.013	-.015	.024[a]	.011
Ireland	-.006	-.014	-.017[b]	-.024[a]	.032[a]	.014[a]
Italy	-.007	-.010			.048[a]	.002
Japan	.000	-.000	-.033	-.032	.023[b]	.011
Netherlands	-.004	-.012			.008	.003
New Zealand	.003	-.007			.013[b]	.004
Norway	.013	.007			.034[a]	.005
Portugal	.013	.008			.037[b]	.005
Spain	-.004	-.018			.010	.007
Sweden	.004	-.004			.021[a]	.005
Switzerland	.001	-.004			.034[a]	-.051
Turkey	-.037[a]	-.039[a]			-.002	.020
United Kingdom	.004	-.008	-.008	-.017	.029[a]	-.000
United States	.017[a]	.014[b]	.021[a]	.017[b]	.020[a]	-.000

Notes. The sample period is 1963-1993. The superscripts a and b indicate statistical significance of the .01 and .05 levels, respectively. Newey-West corrections with a window length of five years were used in assessing statistical significance.

Table 6. Augmented Levin-Lin Statistics for the OECD Sample

Lag Length (q)	*(i)*	*(ii)*	*(iii)*	*(iv)*	*(v)*	*(vi)*
1	-5.79	-6.26	-3.84	-3.81	-5.83	-2.94
2	-5.73	-5.44	-4.05	-4.04	-5.72	-4.01
3	-5.64	-5.57	-4.29	-4.30	-5.30	-5.63

Notes. The sample period extends from 1962+*q* to 1993. Country- and time-specific effects are included in all regressions. The .05 and .01 critical values of the statistics are -7.74 and -8.27 for series (i), (ii), (v), and (vi) and -5.42 and 5.94 for series (iii) and (iv), respectively.

Table 7. Im-Pesaran-Shin Statistics for the OECD Sample

Lag Length (q)	*(i)*	*(ii)*	*(iii)*	*(iv)*	*(v)*	*(vi)*
1	-0.27	-0.50[a]	-0.10	-0.22	0.30	-0.45[b]
2	0.06	0.02	0.17	0.07	0.40	-0.40[b]
3	0.04	0.11	0.15	0.19	0.34	-0.59[a]

Notes. The sample period extends from 1962+*q* to 1993. The statistics for series (i), (ii), (v) and (vi) are asymptotically distributed as N(0,1/22) and the statistics for series (iii) and (iv) are asymptotically distributed as N(0,1/7). The superscripts a and b indicate statistical significance on a one-sided test at the .01 and .05 levels, respectively.

Table 8. Estimates of β

for Four Labor Supply Measures and 22 OECD Countries

Countries	***(i)***	***(ii)***	***(iii)***	***(iv)***
Australia	-.304	-.877[b]		
Austria	.127[b]	.167[b]		
Belgium	.096	.292[a]		
Canada	.439[b]	.283[b]	.629[a]	.221[a]
Denmark	.901[a]	.318[a]		
Finland	.325[a]	.282[a]		
France	.182[a]	.374[a]		
Germany	.098[b]	.221[a]	.291[a]	.350[a]
Greece	.635[a]	.874[a]	.053	-.047
Ireland	.099	.124	.012	.020
Italy	-.022	-.053		
Japan	.561[a]	.544[a]	.375[a]	.354[a]
Netherlands	.563[a]	.747[a]		
New Zealand	.096	.762[a]		
Norway	.238[a]	.181[a]		
Portugal	.524[a]	.332		
Spain	.737[a]	.998[a]		
Sweden	-.066	-.104		
Switzerland	.238[a]	.256[a]		
Turkey	1.318	.338		
United Kingdom	-.163[b]	-.195[a]	-.153[a]	-.198[a]
United States	.194[a]	.172[a]	.214[a]	.136[a]
Average	.310[a]	.274[a]	.203[a]	.119[a]

Notes. The sample period extends from 1967-1991. The subscripts a and b indicate statistical significance at the .001 and .05 levels, respectively. Inference on the averages in the last row is based on the assumption that estimates of β in the previous rows are independently distributed.

Table 9. Estimates of δ_2 for Several PWT Samples

	Full Sample		ICP Sample	
Span	***N***	δ_2	***N***	δ_2
1950-1960	60	.058 (.066)	51	.047 (.072)
1950-1970	60	.041 (.068)	51	.038 (.076)
1950-1980	60	.054 (.058)	51	.027 (.066)
1950-1990	57	.020 (.049)	50	.021 (.059)
1960-1970	125	.062[b] (.029)	79	.022 (.042)
1960-1980	125	.149[a] (.045)	79	.088 (.061)
1960-1990	104	.170[a] (.053)	72	.124[b] (.062)
1970-1980	133	.123[a] (.044)	82	.084 (.062)
1970-1990	108	.180[a] (.058)	75	.196[a] (.060)
1980-1990	111	.036 (.056)	75	.038 (.069)

Notes. *N* is the number of countries in the sample. Standard errors appear in parentheses below their estimates. The superscripts a and b indicate statistical significance at the .01 and .05 significance levels, respectively.

Figure 1. Hours Supplied as Labor in the United States, 1890-1995

Figure 2. Employment as a Percent of the 15-64 Population in 22 OECD Countries, 1960-1994