DISCUSSION PAPER SERIES

No. 5801
CEPR/EABCN No. 32/2006

DRIFT AND BREAKS IN LABOUR PRODUCTIVITY

Luca Benati

INTERNATIONAL MACROECONOMICS

€ABCN
Euro Area Business Cycle Network
www.eabcn.org

Centre for Economic Policy Research
www.cepr.org

Available online at: www.cepr.org/pubs/dps/DP5801.asp

ISSN 0265-8003

DRIFT AND BREAKS IN LABOUR PRODUCTIVITY

Luca Benati, Bank of England

Discussion Paper No. 5801
CEPR/EABCN No. 32/2006
August 2006

Centre for Economic Policy Research
90–98 Goswell Rd, London EC1V 7RR, UK
Tel: (44 20) 7878 2900, Fax: (44 20) 7878 2999
Email: cepr@cepr.org, Website: www.cepr.org

This Discussion Paper is issued under the auspices of the Centre's research programme in **INTERNATIONAL MACROECONOMICS**. Any opinions expressed here are those of the author(s) and not those of the Centre for Economic Policy Research. Research disseminated by CEPR may include views on policy, but the Centre itself takes no institutional policy positions.

The Centre for Economic Policy Research was established in 1983 as a private educational charity, to promote independent analysis and public discussion of open economies and the relations among them. It is pluralist and non-partisan, bringing economic research to bear on the analysis of medium- and long-run policy questions. Institutional (core) finance for the Centre has been provided through major grants from the Economic and Social Research Council, under which an ESRC Resource Centre operates within CEPR; the Esmée Fairbairn Charitable Trust; and the Bank of England. These organizations do not give prior review to the Centre's publications, nor do they necessarily endorse the views expressed therein.

These Discussion Papers often represent preliminary or incomplete work, circulated to encourage discussion and comment. Citation and use of such a paper should take account of its provisional character.

Copyright: Luca Benati

CEPR Discussion Paper No. 5801
CEPR/EABCN No. 32/2006

August 2006

ABSTRACT

Drift and Breaks in Labour Productivity*

We use tests for multiple breaks at unknown points in the sample, and the Stock-Watson (1996, 1998) time-varying parameters median-unbiased estimation methodology, to investigate changes in the equilibrium rate of growth of labor productivity–both per hour and per worker–in the United States, the Eurozone Australia, and Japan over the post-WWII era. Results for the U.S. well capture the 'conventional wisdom' of a golden era of high productivity growth, the 1950s and 1960s; a marked deceleration starting from the beginning of the 1970s; and a strong growth resurgence starting from mid-1990s. Interestingly, evidence suggests the 1990s' productivity acceleration to have reached a plateau over the last few years. Results for the Eurozone point towards a marked deceleration since the beginning of the 1980s, with the equilibrium rate of growth of output per hour falling to 0.9% in 2004:4. Results based on Cochrane's variance ratio estimator suggest a non-negligible fraction of the quarter-on-quarter change in labor productivity growth to be permanent. From a technical point of view, we propose a new method for constructing confidence intervals for variance ratio estimates based on spectral bootstrapping. Preliminary Monte Carlo evidence suggests such a method to possess good coverage properties.

JEL Classification:
Keywords: bootstrapping, frequency domain, median-unbiased estimation, Monte Carlo integration, structural break tests, time-varying parameters and variance ratio

Luca Benati
Senior Economist
Bank of England
Threadneedle Street
London
EC2R 8AH
Email: Luca.Benati@bankofengland.co.uk

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=159554

*The publication of this Paper is funded by the Euro Area Business Cycle Network (www.eabcn.org). This Network provides a forum for the better understanding of the euro area business cycle, linking academic researchers and researchers in central banks and other policy institutions involved in the empirical analysis of the euro area business cycle. I wish to thank Charlie Bean, Tim Cogley, John Fernald, Dale Henderson, Lutz Kilian, Stephen Oliner, John Roberts, and Frank Smets for useful discussions, Mark Watson for extremely helpful suggestions, and participants to the CEPR-Euro Area Business Cycle Network conference 'Productivity and the Business Cycle: Evidence for Europe and the U.S.' – in particular my discussant, Attila Ratfai – for comments. Thanks to Ivan Roberts, of the Reserve Bank of Australia, and Ryo Kato, of the Bank of Japan, for kindly providing data for their countries. Usual disclaimers apply. The views expressed in this paper are those of the author, and do not necessarily reflect those of the Bank of England or the Monetary Policy Committee.

Submitted 29 July 2006

1 Introduction

Changes over time in equilibrium productivity growth are of interest to economists for several reasons. First, in the long run productivity is the key underlying determinant of a society's standards of living, and its possible future evolution plays therefore a crucial role in some of the most hotly debated current policy issues, like the future solvency of pension systems. As stressed by Robert Gordon[1] within the context of the debate on the future of Social Security in the United States, for example,

> [t]here has been insufficient attention in public discussions of the Social Security 'crisis' that the official assumptions about future growth by the Social Security Administration are unbelievably pessimistic. [...] [T]hese assumptions are for growth over the next 75 years in real GDP of 1.4 percent, in the labor force of 0.3 percent, and in business productivity of 1.3 percent. [...] [T]he Social Security Administration has an alternative forecast of 2.14 percent growth in real GDP that puts off the 'day of reckoning' until 2072. Potential output growth of 2.9 percent would put off the day of reckoning until the year 2116 [...].

As Gordon makes clear, even seemingly mild differences in the assumptions concerning potential output growh have markedly different implications for the precise date in which U.S. Social Security will become insolvent. In particular, in the light of both a vast literature documenting the U.S. productivity acceleration since the second half of the 1990s, and the results reported in the present work—with trend productivity growth in both the U.S. nonfarm business and the business sectors estimated at 2.7% at the end of 2005—assuming a trend rate of growth of productivity in the business sector of 1.3% appears indeed as unduly pessimistic.[2]

Second, mis-estimation of the true underlying equilibrium productivity growth rate may lead, in principle, to serious policy mistakes. In a series of influential papers,[3] Athanasios Orphanides has argued, for example, that part of the blame for the Great Inflation should be attributed to the FED's inability to detect, in real time, the productivity slowdown of the beginning of the 1970s, thus leading to an over-estimation of the authentic amount of slack existing in the economy.[4] A conceptually equivalent way of making the same point is that, as stressed by, e.g., Laubach and Williams (2003), changes in the rate of growth of potential output are closely linked to changes in the Wicksellian rate of interest, so that failure to identify

[1] See Gordon (1999).

[2] This does *not* imply, however, that a 'conservative' estimate should be regarded as irrational or unjustified, when risk considerations are taken into account. As this paper shows, indeed, U.S. trend productivity growth has fluctuated quite substantially over the post-WWII era, so that, on strictly logical grounds, a future productivity slowdown should not be ruled out.

[3] See e.g. Orphanides (2003).

[4] For an analysis of the consequences of learning about changes in trend productivity growth within the context of a DSGE model, see Edge, Laubach, and Williams (2004).

shifts in equilibrium productivity growth automatically leads to a mis-estimation of the natural rate of interest, with potentially dire consequences for monetary policy.[5]

In this paper we use tests for multiple breaks at unknown points in the sample, and the Stock-Watson (1996, 1998) time-varying parameters median-unbiased estimation methodology, to investigate changes in the equilibrium rate of growth of labor productivity—both per hour and per worker—in the United States, the Eurozone, Australia, and Japan over the post-WWII period. Based on either the Bai and Perron (1998, 2003) methodology, or the Bai (1997b) method of estimating multiple breaks sequentially, one at a time, structural break tests produce, overall, surprisingly little evidence of time-variation in trend productivity growth. As we argue based on Monte Carlo simulations of estimated time-varying parameters models—and in line with conceptually similar evidence produced by Cogley and Sargent (2005)—the most likely explanation for break tests' failure to detect much evidence of time-variation is that historical changes in equilibrium productivity growth have simply been too gradual to be detectable *via* such a powerful, but ultimately quite crude methodology.

The more flexible Stock and Watson's (1996, 1998) time-varying parameters median-unbiased estimation (henceforth, TVP-MUB) methodology—based on the notion that the underlying DGP may be characterised by random-walk drift—detects indeed strong evidence of time-variation for both the United States and the Eurozone. Results for output per hour in the United States well capture the 'conventional wisdom' of a golden era of high productivity growth, the 1950s and 1960s; a marked deceleration starting from the beginning of the 1970s; and a strong growth resurgence starting from mid-1990s. Interestingly, evidence clearly suggests the 1990s' productivity acceleration to have reached a *plateau* over the last few years. Results for the Eurozone point towards a marked deceleration since the beginning of the 1980s, with the equilibrium rate of growth of output per hour falling to 0.9% in 2004:4. As for the other contries, we detect some mild evidence of time-variation for Japan, while for Australia we identify time-variation in the rate of growth of output per hour, but not in that of output per worker.

Evidence of random-walk time-variation in equilibrium labor productivity growth then naturally induces us to investigate the conceptually related issue of what fraction of the quarter-on-quarter change in labor productivity growth should be regarded as permanent, which we do *via* Cochrane's (1988) variance ratio estimator. We estimate a 90% confidence interval for the the 'size of the unit root' in U.S. labor productivity growth in the non-farm business and in the business sectors to be [2.3; 8.3] and [1.8; 6.4], respectively, with median point estimates equal to 4.0% and 2.8%, thus clearly pointing towards a non-negligible fraction of quarter-on-quarter changes in labor productivity growth as being permanent.

Finally, from a technical point of view, the paper proposes a new method for constructing confidence intervals for variance ratio estimates based on spectral boot-

[5] Conceptually in line with the present work, Laubach and Williams (2003) identify significant changes in the Wicksellian rate of interest in the United States over the post-WWII era.

strapping. Preliminary Monte Carlo evidence suggests such a method to possess good coverage properties.

The immediate implication of our findings for monetary policy is that the problems arising from changes in equilibrium productivity growth discussed by Orphanides within the context of the U.S. Great Inflation of the 1970s should be regarded as part of the normal 'macroeconomic landscape', i.e. of the normal set of problems central banks have to worry about—at least, in the United States.

The paper is organised as follows. The next section presents results from tests for multiple structural breaks at unknown points in the sample in the mean, based on (a) Andrews' (1993) *sup*-Wald and Andrews and Ploberger's (1994) *exp*-Wald test statistics, and Bai's (1997a) method of estimating multiple breaks sequentially, one at a time, and (b) the multiple break tests methodology introduced by Bai and Perron (1998, 2003). In Section 3 we present results based on Stock and Watson's (1996, 1998) TVP-MUB estimation methodology applied to univariate autoregressions for labor productivity growth. Section 4 presents a Monte Carlo investigation of the power of structural break tests conditional on taking, as data generation processes, some of the models estimated in section 3. In section 4 we perform an exercise in the spirit of Edge, Laubach, and Williams (2004), by recursively applying the Stock-Watson methodology to the rates of growth of U.S. output per hour in the nonfarm business and in the business sectors since the beginning of the 1970s, thus computing pseudo real-time estimates of long-run productivity growth. Finally, having detected, in section 3, evidence of random walk time-variation for many series in our dataset, we estimate, in section 5, the size of the permanent component of the quarter-on-quarter change in labor productivity growth, based on Cohrane's (1988) variance ratio estimator. Section 6 concludes.

2 Testing for Breaks in the Mean

2.1 Results based on Bai (1997b), and Andrews (1993) and Andrews and Ploberger (1994)

We start by testing for multiple structural breaks at unknown points in the sample in the mean of labor productivity growth. Our first approach combines the Bai and Perron (2003) method of testing for breaks in the mean by regressing the series on a constant, using the Newey and West (1987) covariance matrix estimator to control for autocorrelation and/or heteroskedasticity in the residuals;[6] the Andrews (1993) and Andrews and Ploberger (1994) *sup*-Wald and *exp*-Wald test statistics; and the Bai (1997a) method of estimating multiple breaks sequentially, one at a time.[7] We

[6] For an application of this methodology to inflation rates and real interest rates, see Rapach and Wohar (2003).

[7] As discussed in Bai (1997a), sequential estimation of the break dates, compared to the alternative simultaneous estimation, presents two key advantages. First, computational savings. Second,

impose 15% symmetric trimming, and we bootstrap the critical values as in Diebold and Chen (1996), setting the number of bootstrap replications to 1,000. Finally, we compute confidence intervals for estimated break dates according to Bai (1997b).

Table 1 reports the results based on the Andrews-Ploberger *exp*-Wald statistic.[8] Quite surprisingly, we identify break dates for only four series, output per hour in the U.S. business and manufacturing sectors, and real GDP per hour and per worker in the Eurozone. Results for the Eurozone well accord with the conventional wisdom notion of a significant productivity slowdown over the most recent period, with the equilibrium (mean) rate of growth of real GDP per hour estimated to have fallen from 2.3% before 1995:1 to 0.9% over the most recent sub-period; and real GDP per worker having even experienced two breaks since the beginning of the 1970s, with equilibrium growth estimated to have decreased from the 3.6% of the first sub-period to a paltry 0.6% over the most recent one.

In the light of the vast literature documenting the U.S. productivity resurgence since mid-1990s, results for the United States appear instead as puzzling. First, although we estimate a break in mean productivity growth for the overall manufacturing sector, with a marked increase from 3.2% before 2001:4 to 5.8% after that, the estimated break date appears clearly at odds with the conventional wisdom notion that the U.S. productivity acceleration started around mid-1990s.[9] Second, as for the business sector our results, with a single estimated break in 1966:2, and a fall in mean productivity growth from 3.5% to 2.1%, uniquely capture the 1970s productivity slowdown, and entirely fail to capture the productivity resurgence of the most recent period. Third, in the light of the previously discussed vast literature documenting the U.S. productivity resurgence of the second half of the 1990s, failure to identify break dates for all other series—in particular, output per hour in the nonfarm business sector, traditionally regarded as the 'bellweather series' for U.S. productivity studies—is especially puzzling. What can account for these results?

There are two possible explanations for our failure to identify much evidence of time-variation in labor productivity growth, which we explore in the next paragraph and in section 4, respectively. The first explanation has to do with a well-known weakness of Bai's (1997a) sequential procedure for break dates estimation when the parameter whose constancy is being tested experiences first a decrease (increase) and then an increase (decrease). Given that in these cases it is comparatively hard to identify the *first* break to begin with, the entire procedure tends to break down, and no break date ends up being estimated. The procedure proposed by Bai and Perron (1998, 2003), which is based on the notion of assessing which, among a set of models with and without breaks, is more likely to have generated the data, is on the other

robustness to misspecification in the number of breaks.

[8] Results based on Andrews' *sup*-Wald test statistics are identical, and therefore are not reported here, but are available from the author upon request.

[9] As we will see in section 3, results based on the Stock-Watson methodology clearly suggest that the productivity acceleration in the U.S. manufacturing sector started in the *first half* of the 1990s.

hand in principle immune to this kind of problem. While in principle this appears, at least for the United States, a likely explanation for our failure to identify much evidence of time-variation, as we will see in the next paragraph *this is clearly not the case*, as results based on the Bai-Perron procedure are near-identical to the ones we just discussed.

The second explanation is that historical changes in equilibrium rates of labor productivity growth may have simply been *too gradual* to be detectable *via* a powerful but, ultimately, intrinsically quite crude procedure such as structural break tests. The evidence produced by Cogley and Sargent (2005) of a sometimes remarkably low power of structural break tests, conditional on taking their estimated Bayesian time-varying parameters VAR as data generation process (henceforth, DGP), provides *prima facie* evidence that this may be the correct explanation. As we will see in section 4, this appears indeed to be the case: conditional on taking as DGPs some of the models estimated *via* the Stock-Watson TVP-MUB estimation methodology, which is characterised by random-walk time-variation, break tests—specifically, both the Andrews-Ploberger *exp*-Wald statistic and the Bai-Perron 'double maximum' test statistics—exhibit a power ranging between 31% and 43%.

2.2 Results based on Bai and Perron (1998, 2003)

Let's now turn to the results from the Bai-Perron methodology. In what follows we exactly follow the recommendations of Bai and Perron (2003),[10] with the only difference that, instead of relying on the asymptotic critical values tabulated in Bai and Perron (1998), and encoded in Pierre Perron's *Gauss* code, we bootstrap both critical and p-values as in Diebold and Chen (1996), setting the number of bootstrap replications to 1,000. We start by looking at the *UDmax* and *WDmax* double maximum test statistics. Conditional on both statistics being significant at the 10% level—thus indicating the presence of at least one break—we decide on the number of breaks by sequentially examining the *sup*-$F(\ell+1|\ell)$ test statistics, starting from the *sup*-$F(2|1)$ one. Finally, we set the maximum allowed number of structural changes to m=4.

Tables 2-4 report the results. As the tables clearly show, compared with the results we saw in the previous paragraph the *only* significant difference concerns the series for U.S. output per hour in the business sector, for which, in line with Fernald (2005), we identify two break dates, in 1973:1 and 1997:1, and a U-shaped evolution of mean productivity growth, first falling from 3.3% to 1.7%, and then increasing to 3.1% over the most recent sub-period.[11] As for the other series, first, as in the previous paragraph, we only identify break dates for U.S. output per hour in the overall manufacturing sector, and for real GDP per hour and per worker in the Eurozone; and second, in all cases the difference between the break dates reported in Table 3 and those reported in Table 1 is at most one quarter.

[10] See Bai and Perron (2003) section 5.5, 'Summary and Practical Recommendations'.

[11] Output per hour in the business sector is the only series considered by Fernald (2005).

Overall, failure to identify much evidence of time-variation in mean labor productivity growth based on structural break tests appears therefore as an extremely robust finding. As we mentioned in the previous paragraph, however, one possible explanation for our results is that historical changes in equilibrium labor productivity growth may have been so gradual as to be hard to detect *via* break tests. A simple way of formalising, from an econometric point of view, the notion of 'gradual change' in the underlying DGP is *via* time-varying parameters models, and we therefore now turn to the Stock and Watson (1996, 1998) TVP-MUB methodology, which presents the attractiveness of allowing the researcher to test for the presence of random-walk time-variation in the data, against the null of time-invariance, and then to estimate its extent.

3 Estimating Models of Random-Walk Time-Variation

In this section we present results based on the Stock and Watson (1996) and Stock and Watson (1998) TVP-MUB methodology applied to the AR(p) model[12]

$$y_t = \mu + \phi_1 y_{t-1} + \phi_2 y_{t-2} + ... + \phi_p y_{t-p} + u_t \tag{1}$$

where y_t is the rate of growth of labor productivity. We select the lag order, p, based on the Bayes information criterion, for a maximum possible number of lags P=6. With a single exception discussed below, concerning the issue of how to tackle the possible presence of heteroskedasticity in the data—for which we adopt a solution along the lines of Boivin (2004)—we closely follow Stock and Watson (1996).

Letting θ_t=[μ_t, $\phi_{1,t}$, ..., $\phi_{p,t}$]$'$ and z_t=[1, $y_{t-1,t}$, ..., $y_{t-p,t}$]$'$, the time-varying parameters version of (1) is given by:

$$y_t = \theta_t' z_t + u_t \tag{2}$$

$$\theta_t = \theta_{t-1} + \eta_t \tag{3}$$

with η_t *iid* $N(0_{p+1},\ \lambda^2\sigma^2 Q)$, with 0_{p+1} being a (p+1)-dimensional vector of zeros; σ^2 being the variance of u_t; Q being a covariance matrix; and $E[\eta_t u_t]$=0. Following

[12] Our choice to work with (time-varying) univariate autoregressions deserves some discussion. In principle, we could have chosen to work with more sophisticated models along the lines of (eg) Roberts (2001). In practice, however, there are two issues to take into account. First, in order to investigate (time-variation in) *simple features of a time series*—like its persistence, volatility, or, in the present case, its mean (equilibrium level)—*sophisticated models are not needed.* The starkest possible illustration of this is the fact that the Great Stability in the United States was first robustly identified by Kim and Nelson (1999) and McConnell and Perez-Quiros (2000) *based on univariate methods.* A second key justification for our preference for what we would label as a 'minimalist econometric approach' has to do with the fact that, the more complex the model, the greater the number of possibly questionable assumptions necessarily becomes. Given that, in order to robustly identify simple stylised facts, complicated models are not needed, it is not clear (at least, to us) why unnecessary risks should be taken.

Nyblom (1989) and Stock and Watson (1996, 1998), we set Q=$[E(z_t z_t')]^{-1}$. Under such a normalisation, the coefficients on the transformed regressors, $[E(z_t z_t')]^{-1/2} z_t$, evolve according to a (p+1)-dimensional standard random walk, with λ^2 being the ratio between the variance of each 'transformed innovation' and the variance of u_t.[13]

3.1 Searching for random-walk time-variation

Our point of departure is the Hansen (1999) grid bootstrap median-unbiased estimate of the sum of the autoregressive coefficients, ρ, in (1)—for details, see Hansen (1999). Conditional on the grid bootstrap estimate of ρ we estimate model (1); we compute the residuals, $\hat{u}_t$, and the estimate of the innovation variance, $\hat{\sigma}^2$; and we perform an *exp*- or *sup*-Wald joint test for a single break at an unknown point the sample in μ and ρ, using the Newey and West (1987) covariance matrix estimator to control for possible autocorrelation and/or heteroskedasticity in the residuals. We estimate the matrix Q as in Stock and Watson (1996) as

$$\hat{Q} = \left[T^{-1} \sum_{t=1}^{T} z_t z_t' \right]^{-1}. \tag{4}$$

We start by considering a 30-point grid of values for λ over the interval [0, 0.05], which we call Λ. For each $\lambda_j \in \Lambda$ we compute the corresponding estimate of the covariance matrix of η_t as $\hat{Q}_j$=$\lambda_j^2 \hat{\sigma}^2 \hat{Q}$, and conditional on $\hat{Q}_j$ we simulate model (2)-(3) 10,000 times as in Stock and Watson (1996, section 2.4), drawing the pseudo innovations from pseudo random *iid* $N(0,\ \hat{\sigma}^2)$. For each simulation, we compute an *exp*- or *sup*-Wald test—without however applying the Newey and West (1987) correction, obviously—thus building up its empirical distribution conditional on λ_j. Based on the empirical distributions of the test statistic we then compute the median-unbiased estimate of λ as that particular value of λ_j which is closest to the statistic we previously computed based on the actual data. In case the *exp*- or *sup*-Wald test statistics computed based on the actual data are greater than the corresponding medians of the empirical distributions conditional on λ_j=0.05, we add one more step to the grid, and we estimate λ as 0.05172. Finally, we compute the p-value based on the empirical distribution of the test conditional on λ_j=0.

Table 5 reports the results. Starting from the United States we detect, based on the simulated p-values, evidence of time-variation at the 10% level for all the series for output per hour, with the single exception of the nondurables manufacturing sector (for which, however, MUB estimates of λ are comparatively large based on either test statistic). Evidence is especially strong, first—and not surprisingly, in the light of the results reported in section 2—for the overall manufacturing sector; and

[13] To be precise, given that the Stock-Watson methodology is based on local-to-unity asymptotics, λ is actually equal to the ratio between τ, a small number which is fixed in each sample, and T, the sample length.

second, for the nonfarm business sector, with simulated p-values based on the two statistics of 0.023 and 0.006, respectively. Results for the output per worker series are, on the other hand, mixed, with strong evidence of time-variation for the overall manufacturing sector and for durables production, mixed evidence for non-durables, and weak evidence for the other sectors. Turning to other countries, for Japan and the Eurozone we detect weak and very strong evidence, respectively, of time-variation, and for Australia we identify strong evidence for output per hour in all industries, weaker evidence—quite surprisingly—for output per hour in the market sector, and no evidence at all for real GDP per worker, with a MUB estimate of λ being exactly zero.

3.2 Estimating time-varying equilibrium productivity growth

We now proceed to compute time-varying estimates of equilibrium productivity growth rates and, crucially, confidence bands around the estimates. We take into account of both filter and parameter uncertainty *via* the modification for the problem at hand of the Hamilton (1986)[14] Monte Carlo integration procedure described in Appendix C. To this purpose, a necessary preliminary step is deconvoluting the probability density function of $\hat{\lambda}$, which we do *via* the procedure described in Appendix B. Figures 3 and 4 show the deconvoluted PDFs of $\hat{\lambda}$, together with the corresponding MUB estimates of λ, for all the series for which the MUB estimate is greater than zero.

In what follows we present results for all the series for which the MUB estimate of λ is greater than zero, disregarding therefore the fact that the simulated p-value is, or is not, smaller than 10%. The key reason for doing so is that a p-value above 10% should be regarded as significant evidence against time-variation if and only if the researcher had very compelling reasons for believing in time-invariance. It is not clear at all, however, why this should be the case—to put it differently, it is not clear why the hypothesis of time-invariance should be granted such a privileged status—and in what follows we therefore report results for all series for which empirical evidence does not manifestly point towards time-invariance.[15]

Before proceeding further, it is necessary to briefly discuss the main difference betwen the approach adopted herein and the one found in, e.g., Stock and Watson (1996), concerning how we tackle the possible presence of heteroskedasticity in the data. As stressed by Stock (2002) in his discussion of Cogley and Sargent (2002),[16] estimating time-varying parameters models without controlling for the possible pres-

[14] See also Hamilton (1985).

[15] A second reason is that the MUB *point* estimate and the simulated p-value should not be regarded as the only relevant pieces of information. Consider, for example, the case of output per worker in the U.S. business and nonfinancial corporations sectors. Although the p-values reported in Table 5 are above 36% for both series, as Figure 4 makes clear a significant fraction of the probability mass of $\hat{\lambda}$ corresponds to comparatively large values of λ. Finally, the pseudo real-time experiment of section 5 suggests that, in general, p-values possess limited informational content.

[16] See also Cogley (2005).

ence of heteroskedasticity causes a systematic overestimation of the authentic extent of coefficients' drift, as the imposition of a constant covariance structure forces the time-varying parameters to 'pick up' part of the variation in the data originating from time-variation in the covariance. In what follows we adopt a solution along the lines of Boivin (2004), testing for multiple structural breaks at unknown points in the sample in the innovation variance in equation (1),[17] based on either the *exp*- or the *sup*-Wald test statistics, and the Bai (1997a) method of estimating multiple breaks sequentially, one at a time,[18] bootstrapping the critical values as in Diebold and Chen (1996), and imposing 15% symmetric trimming. Finally, we compute confidence intervals for estimated break dates as in Bai (1997b). Results based on the *exp*-Wald statistic are shown in Table 3:[19] we detect volatility breaks for only five series, output per hour and per worker in the U.S. nonfarm business and business sectors, and real GDP per hour in the Eurozone.

Based on the median-unbiased estimates of λ, on the deconvoluted PDFs of $\hat{\lambda}$, and on the estimated breaks in the innovation variance, we then estimate time-varying equilibrium rates of labor productivity growth, and confidence bands around the estimates, by taking into account of both parameter and filter uncertainty *via* the Monte Carlo integration procedure described in Appendix C. Figure 5-7 show the results.

3.2.1 Evidence for the United States

Starting from output per hour (Figure 5), results for the nonfarm business sector—traditionally regarded as the 'bellweather series' for U.S. productivity studies—well accord with conventional wisdom, pointing towards

- a former golden era of comparatively high productivity growth, until the first half of the 1960s, with our preferred measure of trend growth—the median of the distribution of $\gamma_{t|T} \equiv \mu_{t|T}/(1\text{-}\phi_{1,t|T}\text{-}...\text{-}\phi_{p,t|T}) \equiv \mu_{t|T}/(1\text{-}\rho_{t|T})$, computed based on the deconvoluted PDF of $\hat{\lambda}$—[20]estimated between 2.5% and 2.7%;

[17]As stressed by Boivin (2004, footnote 16), the estimation of different variances for different sub-samples is indeed 'entirely consistent with the TVP specification, asymptotically', given the assumption of local-to-zero time variation. Although Boivin (2004) considered a single break—estimating two different variances for the pre-Volcker and post-1979 periods—his approach is entirely appropriate also in the case of multiple breaks. (I wish to thank Mark Watson for confirming this to me.)

[18]An alternative would have been to adopt a Bayesian approach, which would have allowed us to model time-variation in the innovation variance *via* a stochastic volatility model along the lines of, e.g., Jacquier, Polson, and Rossi (2004). We have preferred to adopt the present, Classical approach as the adoption of a Bayesian perspective would have compelled us to specify a prior for the extent of random-walk drift, which we want instead to entirely estimate from the data.

[19]Results based on the *sup*-Wald statistic are identical, and are not reported here, but are available from the author upon request.

[20]The reason why this is our preferred estimate is because—different from the one conditional on

- a marked slowdown from the first half of the 1960s up to around 1980, with equilibrium productivity growth estimated to have fallen to 1.8% in 1980:2;
- a period of stagnation, the 1980s, with trend growth fluctuating between 1.8% and 2.0%; and
- a growth resurgence—first tentative, and then, since mid-1990s, literally explosive—starting from the beginning of the 1990s, with equilibrium growth estimated, for the latest quarter of our sample, 2005:4, at 2.7%. Interestingly, evidence clearly points towards the productivity acceleration to have reached a *plateau* over the last few years.

Quite intriguingly, 2.7% is very close to to the numerical estimate recently obtained for the non-farm business sector by Jorgenson, Ho, and Stiroh (2004), 2.6%, based on a completely different methodology, growth accounting.[21][22]

Results for the business sector are (not surprisingly) broadly in line with those for the nonfarm business sector—in particular, the evidence of a *plateau* in trend growth reached over the most recent years is even slight clearer—with the main difference being the steady and consistent deceleration in trend growth from the beginning of the sample up until around 1980. After fluctuating around 2.3% until the beginning of the 1990s, equilibrium growth is estimated to have strongly accelerated over the following years, reaching, in 2005:4, 2.7%.[23]

Evidence for nonfinancial corporations is qualitatively, although not quantitatively, in line with that discussed so far, with a U-shaped evolution of trend growth over the post-WWII era, the main differences being (*i*) a smaller extent of variation from the beginning of the sample until the end of the 1980s, with the minimum and the maximum of trend growth estimated at 1.7% and 2.3%, respectively; (*ii*) a much stronger acceleration since the beginning of the 1990s, with trend growth increasing from 2.0% in 1990:1 to 3.2% in the last quarter of the sample, 2003:3; and (*iii*) no evidence of a *plateau* reached over the most recent period. Under this respect, however, it is important to stress how, given the shorter sample period, these results

the TVP-MUB estimates of λ—it takes into account of all possible sources of uncertainty. Estimates conditional on the TVP-MUB estimates of λ are however, in general, very close to the median estimates, and are available upon request.

[21] At first sight, our estimates may appear not to be comparable with the Jorgenson-Ho-Stiroh ones, which are *forward-looking*—specifically they are projections for trend productivity growth in the nonfarm busines sector *over the next decade*. It is important to stress, however, that given the nature of the Stock-Watson TVP-MUB method used herein, which is based on the assumption of random-walk time-variation, the most recent estimate is automatically a projection into the infinite future, so that our estimates and the Jorgenson-Ho-Stiroh ones are *exactly* comparable.

[22] Jorgenson *et al.*'s estimates are based on the 2004:2 data vintage. In two previous versions of this work, based on the 2004:3 and 2005:2 data vintages, we estimated equilibrium productivity growth in the non-farm business sector at 2.6% in both cases.

[23] These results are therefore in line with those based on the Bai-Perron methodology discussed in section 2.2, and with Fernald's (2005).

are strictly speaking not incompatible with those for the nonfarm business and the business sectors.

Turning to manufacturing, the main findings emerging from the bottom row of Figure 5 are

- a dramatic productivity acceleration in the overall manufacturing sector starting around 1990, with trend growth increasing from 3.0% in 1990:1 to 4.6% in 2005:4,[24] and clear evidence of a *plateau* reached since about 2003; and
- significant differences between the durables and nondurables goods sectors, with trend growth for nondurables increasing from 2.1% in 1990:1 to 2.7% in 2005:4, and with the corresponding figures for the production of durables being 3.5% and 5.8%, respectively.[25]

Let's now turn to output per worker (Figure 6). Given that the MUB estimate of λ for the nonfarm business sector is exactly zero, the figure only reports evidence for the other five series. For each series, the overall identified pattern of change closely mimics the one for the corresponding series in Figure 5, but, not surprisingly, with significantly less variation. For the business sector, for example, trend growth is estimated to have been equal to 2.6% at the beginning of the sample, to have decreased to a minimum of 2.1% in 1988:4, and to have increased to 2.4% in the last quarter of the sample, 2003:4, with the corresponding figures for output per hour being 3.2%, 2.1%, and 2.7% respectively. Results for the other series exhibit even less time-variation, with the minimum and the maximum of trend growth being 1.7% and 2.3%, respectively, for nonfinancial corporations, and 2.2% and 2.5% for the nondurables manufacturing sector.

Finally, a point worth stressing is that although, up until now, we have uniquely focussed on median *point* estimates, as the figures clearly show, the extent of uncertainty associated with these estimates—once taking into account of both filter and parameter uncertainty—is quite substantial. Focussing on output per hour in the nonfarm business sector, for example, the width, in percentage points, of a 90%-coverage confidence band was equal to 1.6 at the beginning of the sample, it decerased to a minimum of 1.2 around 2000, and it increased, again, to 1.6 at the very end of the sample.

3.2.2 Evidence for other countries

Turning to other countries, results for the Eurozone need no comment—the first column of Figure 7 already speaks volume. Consistent with the 'Eurosclerosis' con-

[24] At first blush 4.6% may appear as implausibly high. It is therefore worth stressing that the *simple average* of the annualised quarter-on-quarter rate of growth of output per hour in the overall manufacturing sector *has been equal, since 2000:1, to 4.9% ...*

[25] Again, 5.8% may appear as absurdly high. Again, it is worth stressing that the *simple average* of the annualised quarter-on-quarter rate of growth of output per hour in the durables manufacturing sector has been equal, since 2000:1, to *6.1%*.

ventional wisdom, results based on either real GDP per hour or real GDP per worker point towards a collapse since the beginning of the 1980s and of the 1970s, respectively. Equilibrium growth of output per hour is estimated to have fallen from 2.3% in 1980:3 to a minimum (so far) of 0.9% in 2004:4, with the corresponding figures for output per worker being 3.0% in 1970:3, and 0.7%, in 2003:4.

Consistent with the comparatively large p-values reported in Table 5, evidence for Japan—either per hour or per worker—consistently points towards very little time-variation over the sample period, with trend GDP per hour fluctuating between a minimum of 1.9% and a maximum of 2.0%, with the corresponding figures for GDP per worker being 1.3% and 1.6%.

Finally, in Australia trend productivity growth has clearly accelerated, since the beginning of the 1980s, in the market sector, with the estimated equilibrium going from 1.6% in 1978:3 to 2.3% in 2004:3. Output per hour in the whole economy, on the other hand, first, appears as much less dynamic, fluctuating between a minimum of 1.3% and a maximum of 2.0%. And second, its trend appears, overall, as much less clear, with growth 'sputtering along' until the beginning of the 1990s, then increasing until about 1997, and slightly decreasing ever since.

4 Why Do Break Tests Identify So Little Evidence of Time-Variation? A Monte Carlo Investigation

Conceptually in line with Cogley and Sargent (2005), in section 2 we conjectured that break tests' failure to identify much evidence of time-variation may originate from the fact that historical changes in equilibrium productivity growth may have been too gradual to be detectable *via* such a comparatively 'crude' methodology.[26] In this section we provide some tentative evidence on the plausibility of this conjecture *via* the following Monte Carlo experiment.

We consider three series for which break tests *did not* detect evidence of time-variation, but for which the TVP-MUB methodology *did* identify random-walk time-variation at (at least) the 10% level based on either the *exp*- or the *sup*-Wald test statistics. The three series are output per hour in the U.S. nonfarm business and nonfinancial corporations sectors, and Australia's output per hour in all industries. Based on the DGPs we estimated in section 3 *via* the TVP-MUB methodology we then generate, for each series, 1,000 artificial samples of length equal to the sample length of the corresponding actual series,[27] and for each simulation (i) we perform

[26] Cogley and Sargent (2005) report the following values for the power of the test for the equations for the nominal rate, unemployment, and inflation in their Bayesian time-varying parameters VAR. Andrews (1993)'s *sup*-LM test: 0.136, 0.172, and 0.112. Nyblom (1989)-Hansen (1992) test: 0.076, 0.170, 0.086. Andrews (1993)'s *sup*-Wald test: 0.173, 0.269, 0.711.

[27] Actually, letting T be the sample length of the actual series, we generate artificial samples of length T+100, and we then discard the first 100 observations.

an Andrews-Ploberger *exp*-Wald test for a single break at an unknown point in the sample in the mean (i.e., exactly the same break test we performed in section 2.1), and (*ii*) we test for multiple breaks in the mean based on the Bai-Perron methodology. In both cases we bootstrap the critical values as in Diebold and Chen (1996), setting the number of bootstrap replications to 1,000, exactly as in section 2.

Table 7 reports, for each series, the fraction of times that the null of time–invariance gets rejected based on the *exp*-Wald and, respectively, Bai and Perron's *WDmax* test statistics.[28] As the table makes clear, state-of-the-art break tests fail to reject the (incorrect) null of no time-variation a significant fraction of the times. In the case of the U.S. nonfarm business sector, for example, time-invariance is rejected less than two times out of five based on the *WDmax* statistic, and even less, one time out of three, based on the *exp*-Wald statistic. Although these results ought necessarily to be regarded as preliminary, taken together with those produced by Cogley and Sargent (2005) they provide tentative evidence that our conjecture may indeed be correct. If that's the case, a necessary corollary is that, in order to investigate time-variation in labor productivity growth, break tests are *not* the way to go, and that models of random-walk time variation may provide a more appropriate description of reality.

5 Back to the Future: Computing Pseudo-Real-Time Estimates of Equilibrium Productivity Growth

Suppose that the Great Inflation truly resulted from the Fed's inability to detect, in real time, the productivity slowdown, and further assume that James Stock and Mark Watson had been magically catapulted into the Federal Reserve Board during the first half of the 1970s.[29] Would they have been able, by applying the TVP-MUB methodology, to save the day?

Figure 8 provides some (admittedly, extremely tentative) evidence on this, by showing results from *recursively* applying the Stock-Watson's TVP-MUB methodology to the rates of growth of output per hour in the U.S. nonfarm business and business sectors, for every quarter out of four starting from 1970:4. The methodology we apply is exactly the same we discussed in section 3: for every recursive sample (*i*) we compute the MUB estimate of λ based on the *exp*-Wald test statistic, exactly as in section 3.1; (*ii*) we perform multiple break tests in the innovation variance as described in section 3.2; and (*iii*) we compute median estimates of trend growth, and 90% confidence bands, as in section 3.2. One obvious limitation of the present experiment is that, being based on *revised* data, is only *pseudo* real-time. Unfortunately, the real-time dataset used in Edge, Laubach, and Williams (2004), generously provided to us by John Williams, only contains *annual* data, making it impossible

[28] Results based on the *UDmax* statistic are near-identical, and are available upon request.

[29] Maybe, in a plutonium-powered DeLorean.

to apply to it a 'data hungry' methodology like the TVP-MUB one. Unwillingly, we have therefore decided to perform our experiment based on revised data.

For either series, the bottom row of figure 8 shows, for each quarter out of four (1) the pseudo real-time median estimate of equilibrium productivity growth, together with the 90% confidence bands, computed conditional on the recursive sample ending in that quarter (the thicker lines);[30] and (2) the median estimates, together with the 90% confidence bands, conditional on the full samples up to 2005:4, i.e., exactly the same objects plotted in the first two panels of the top row of Figure 5 (the thinner lines). The top row of Figure 8, on the other hand, shows, for either series, and for each quarter out of four, recursive MUB estimates of λ, and recursively computed simulated p-values for the null hypothesis of no time-variation.

Quite strikingly, the evidence reported in Figure 8 suggests that the recursive application of the TVP-MUB methodology starting from the beginning of the 1970s would have most likely failed to detect the productivity slowdown in real time. In particular, the figure clearly shows how

- simulated p-values stay quite remarkably high until the beginning of the 1980s, thus providing further evidence on their comparatively low informational content. (This provides a further justification for our decision, in section 3.2, to report results for all series for which the MUB estimate of λ is strictly greater than zero.)

- For the nonfarm business sector, the MUB estimate of λ is consistently equal to zero until the very end of the 1970s, while for the business sector it is equal to zero about one-third of the times.

- Crucially, the bottom panels show how, assuming the thin lines to represent the 'truth',[31] for either series pseudo real-time estimates would have consistently missed it until the very beginning of the 1980s. As the figure shows, the extent of over-estimation of trend productivity growth during the 1970s is definitely not negligible, with, in the case of the business sector, pseudo real-time median estimates even breaching, several times, the 'true' 90% upper band.

Results for the period since the beginning of the 1980s do not provide much reassurance on the ability of the best available econometric techniques to provide reliable aid to policymakers, with the recursive trend growth estimates staying, most of the times, quite far away from those conditional on the entire sample. The productivity

[30] In order to avoid confusion, the recursively computed 90% confidence *interval* for quarter t is computed by first getting the confidence *bands* for the recursive sample ending at t (*via* the Monte Carlo integration procedure we already discussed), and then simply by taking, from these bands, the last observation, i.e. the one corresponding to quarter t. We compute these confidence bands *via* Monte Carlo integration even in the case in which the MUB estimate of λ is exactly zero, by drawing from the deconvoluted PDF as described in appendix C.

[31] Needless to say, quite a leap of faith.

resurgence of the 1990s, in particular, appears as especially intriguing, as it clearly represents, for both series, a mirror image of the productivity slowdown of the 1970s. Exactly as in the 1970s the TVP-MUB methodology consistently overestimated the authentic trend growth rate, during the 1990s it consistently underestimated it, catching up with reality only at the very end of the sample.

There is no need to repeat, once again, the obvious limitations of the present exercise—in particular, its pseudo real-time nature. Even with these limitations, however, we believe that these results suggest the need to use some caution in applying even the very best available econometric techniques to policymaking.

6 How Large Is the Permanent Component of Labor Productivity Growth?

Having detected evidence of random-walk time-variation for several series, a natural question to ask is then: 'How large is the permanent component of labor productivity growth?' To put it differently, what fraction of the quarter-on-quarter change in the rate of growth of labor productivity should be regarded as permanent? In order to answer this question, in this section we present results based on Cochrane (1988)'s variance ratio

$$V_k = k^{-1} \frac{\mathrm{Var}\,(y_t - y_{t-k})}{\mathrm{Var}\,(y_t - y_{t-1})} \frac{T}{T-k+1} \tag{5}$$

which we estimate *via*

$$\hat{V}_k = \frac{T}{T-k+1} \left[1 + 2 \sum_{j=1}^{k-1} \frac{k-j}{k} \hat{\rho}_j \right] \tag{6}$$

where the $\hat{\rho}_j$'s are the sample autocorrelations of the first difference of y_t. We construct confidence intervals for $\hat{V}_k$ via the non-parametric spectral bootstrap procedure described in Appendix D. Given that variance ratio estimators strive to estimate a characteristic—the size of the unit root—pertaining to the infinite long-run of a series, we only consider series with at least 40 years of observations. In practice, this compels us to uniquely focus on the series for output per hour and per worker in the U.S. nonfarm business, business, and non-financial corporations sectors.

Figure 9 shows, for the seven series, the median, the mean, and the 90% upper and lower percentiles of the bootstrapped distributions of $\hat{V}_k^*$ at horizons from 1 quarter to 20 years, while Table 8 reports $\hat{V}_k$, together with the median, the mean, and the 90% upper and lower percentiles of the bootstrapped distribution of $\hat{V}_k^*$, at the 20-year horizon. As the figure makes clear, for all series both the median and the mean of the bootstrapped distribution of $\hat{V}_k^*$ become essentially flat at the 20-year horizon, thus clearly suggesting that variance ratio estimates have 'stabilised'. Results in Table 8 clearly show how, in both countries, a non-negligible fraction of the quarter-on-quarter change in the rate of growth of output either per hour or per worker should

be regarded as permanent, with 90%-coverage confidence intervals ranging between [1.8; 6.3] in the U.S. business sector to [2.8; 9.0] in the U.S. nonfinancial corporations sector. The immediate, obvious implication of these finding for monetary policy is that the problems arising from changes in the rate of productivity growth discussed by Orphanides within the context of the U.S. Great Inflation of the 1970s should be regarded as part of the normal 'macroeconomic landscape', i.e. of the normal set of problems central banks have to worry about—at least, in the United States and the United Kingdom.

7 Conclusions

What message(s) should a policymaker take from this paper? Essentially three, we believe.

- First, equilibrium (trend) labour productivity growth should be regarded, *in general*, as time-varying.
- This leads us to a second crucial point. Given that, when changes in trend productivity growth do take place, even the very best available econometric techniques may turn out to be of limited help to policymakers, this naturally suggests the necessity of supplementing such techniques with any possible piece of additional evidence, anecdotal or otherwise.
- When time-variation in equilibrium productivity growth does take place, it is most likely to take place *gradually*—ie without sudden jumps—so that the best way of analysing it is via time-varying parameters models, rather than *via* break tests.

References

Andrews, D. K. (1993): "Tests for Parameter Instability and Structural Change with Unknown Change Point," *Econometrica*, 61, 821–856.

Andrews, D. K., and W. Ploberger (1994): "Optimal Tests When a Nuisance Parameter is Present Only Under the Alternative," *Econometrica*, 62(6), 1383–1414.

Bai, J. (1997a): "Estimating Multiple Breaks One at a Time," *Econometric Theory*, 13(3), 315–352.

——— (1997b): "Estimation of a Change Point in Multiple Regression Models," *Review of Economics and Statistics*, 79, 551–563.

Bai, J., and P. Perron (1998): "Estimating and Testing Linear Models with Multiple Structural Changes.," *Econometrica*, 66(1), 47–78.

——— (2003): "Computation and Analysis of Multiple Structural Change Models," *Journal of Applied Econometrics*, 18(1), 1–22.

Beltrao, K. I., and P. Bloomfield (1987): "Determining the Bandwidth of a Kernel Spectrum Estimate," *Journal of Time Series Analysis*, 8(1), 21–38.

Berkowitz, J., and F. X. Diebold (1998): "Bootstrapping Multivariate Spectra," *Review of Economics and Statistics*, 80(4), 664–666.

Berkowitz, J., and L. Kilian (2000): "Recent Developments in Bootstrapping Time Series," *Econometric Reviews*, 19, 1–48.

Boivin, J. (2004): "Has U.S. Monetary Policy Changed? Evidence from Drifting Coefficients and Real-Time Data," Columbia University, mimeo.

Brillinger, D. R. (1981): *Time Series: Data Analysis and Theory.* New York, McGraw-Hill.

Chow, G., and A. Lin (1971): "Best Linear Unbiased Interpolation, Distribution, and Extrapolation of Time Series by Related Series," *The Review of Economics and Statistics*, 53(4), 372–375.

Cochrane, J. H. (1988): "How Big Is the Random Walk in GNP?," *Journal of Political Economy*, 96(5), 893–920.

Cogley, T. (2005): "How Fast Can the New Economy Grow?," *Journal of Macroeconomics*, 27(2), 179–207.

Cogley, T., and T. J. Sargent (2002): "Evolving Post-WWII U.S. Inflation Dynamics," *in B. Bernanke and K. Rogoff, eds. (2002), NBER Macroeconomics Annuals 2001.*

——— (2005): "Drifts and Volatilities: Monetary Policies and Outcomes in the Post WWII U.S.," *Review of Economic Dynamics*, 8(April), 262–302.

Diebold, F. X., and C. Chen (1996): "Testing Structural Stability with Endogenous Breakpoint: A Size Comparison of Analytic and Bootstrap Procedures," *Journal of Econometrics*, 70(1), 221–241.

Edge, R., T. Laubach, and J. Williams (2004): "Learning and Shifts in Long-Run Productivity Growth," Federal Reserve Board and Federal Reserve Bank of San Francisco, mimeo.

Fernald, J. (2005): "Trend Breaks, Long Run Restrictions, and the Contractionary Effect of Technology Shocks," Federal Reserve Bank of San Francisco, mimeo.

Gordon, R. (1999): "Has the New Economy Rendered the Productivity Slowdown Obsolete?," Northwestern University, mimeo.

Hamilton, J. (1985): "Uncovering Financial Markets Expectations of Inflation," *Journal of Political Economy*, 93(6), 1224–1241.

——— (1986): "A Standard Error for the Estimated State Vector of a State-Space Model," *Journal of Econometrics*, 33(3), 387–397.

Hansen, B. (1992): "Testing for Parameter Instability in Linear Models," *Journal of Policy Modelling*, 14, 517–533.

Hansen, B. (1999): "The Grid Bootstrap and the Autoregressive Model," *Review of Economics and Statistics*, 81(4), 594–607.

Jacquier, E., N. G. Polson, and P. Rossi (2004): "Bayesian Analysis of Stochastic Volatility Models,," *Journal of Business and Economic Statistic*, 12, 371–418.

Jorgenson, D. W., M. S. Ho, and K. J. Stiroh (2004): "Will the U.S. Productivity Resurgence Continue?," *Federal Reserve Bank of New York Economic Policy Review*, 10(13).

Kim, C. J., and C. Nelson (1999): "Has the U.S. Economy Become More Stable? A Bayesian Approach Based on a Markov-Switching Model of the Business-Cycle," *Review of Economics and Statistics*, 81, 608–616.

Laubach, T., and J. Williams (2003): "Measuring the Natural Rate of Interest," *The Review of Economics and Statistics*, 85(4), 1063–1070.

McConnell, M., and G. Perez-Quiros (2000): "Output Fluctuations in the United States: What Has Changed Since the Early 1980s?," *American Economic Review*, 90, 1464–1476.

Newey, W., and K. West (1987): "A Simple Positive-Semi-Definite Heteroscedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 55, 703–708.

Nyblom, J. (1989): "Testing for the Constancy of Parameters Over Time," *Journal of the American Statistical Association*, 84(405), 223–230.

Orphanides, A. (2003): "The Quest for Prosperity Without Inflation," *Journal of Monetary Economics*, 50(April 2003), 633–663.

Rapach, D., and M. Wohar (2003): "Regime Changes in International Real Interest Rates: Are They a Monetary Phenomenon?," *Journal of Money, Credit, and Banking, forthcoming.*

Roberts, J. (2001): "Estimates of Productivity Trend Using Time-Varying Parameter Techniques," *Contributions to Macroeconomics*, 1(1), 1–30.

Stock, J. (2002): "Discussion," *in B. Bernanke and K. Rogoff, eds. (2002), NBER Macroeconomics Annuals 2001.*

Stock, J., and M. Watson (1996): "Evidence of Structural Instability in Macroeconomic Time Series Relations," *Journal of Business ands Economic Statistics*, 14(1), 11–30.

——— (1998): "Median-Unbiased Estimation of Coefficient Variance in a Time-Varying Parameter Model," *Journal of the Americal Statistical Association*, 93(441), 349–358.

A The Data

A.1 United States

Seasonally adjusted quarterly series for output per hour of all persons in the nonfarm business sector (acronym: OPHNFB), in the business sector (acronym: OPHPBS), and in the manufacturing sector (OPHMFG) are from the *U.S. Department of Labor, Bureau of Labor Statistics*. The sample periods are 1987:1-2005:4 for the manufacturing sector, and 1947:1-2005:4 for the other two series. Seasonally adjusted quarterly series for output per hour for nonfinancial corporations, manufacturing durable goods sector, and manufacturing nondurable goods sector are from the *Center for the Study of Innovation and Productivity*'s (henceforth, *CSIP*) website at the Federal Reserve Bank of San Francisco. The sample periods are 1958:1-2003:3 for nonfinancial corporations, and 1987:1-2003:4 for the two other series.

Seasonally adjusted quarterly series for output per worker in the nonfarm business, business, nonfinancial corporations, manufacturing, manufacturing durable goods, and manufacturing nondurable goods sectors are all from the *CSIP*'s website. The sample periods are 1947:1-2003:4 for the nonfarm business and business sectors; 1958:1-2003:3 for the nonfinancial corporations sector; and 1987:1-2003:4 for the three other series.

A.2 Eurozone

A quarterly seasonally adjusted series for real GDP per hour worked has been computed as the ratio between the synthetic euro-area real GDP series from the ECB's Area Wide Model database, and an interpolated quarterly series for overall hours worked in the eurozone, which has been kindly provided by the *European Central Bank*.[32] The sample period is 1980:1-2004:4. A quarterly seasonally adjusted series for real GDP per worker has been computed as the ratio between the synthetic real GDP and employment series from the from the ECB's Area Wide Model database. The sample period is 1970:1-2004:4.

A.3 Australia

Two quarterly seasonally adjusted series for output per hour worked, for the whole economy and the market sector respectively, have been kindly provided by Ivan Roberts of the *Reserve Bank of Australia*. The sample period is 1978:1-2004:3. A quarterly seasonally adjusted series for real GDP per worker has been computed as the ratio between real GDP and employment series from the *IMF*'s *International Financial Statistics*. The sample period is 1982:3-2005:2.

[32] Interpolation has been performed *via* the Chow and Lin (1971) procedure.

A.4 Japan

Seasonally adjusted real GDP per worker, and real GDP per hour worked series have been computed based on data for real GDP, employment, and hours kindly provided by Ryo Kato of the *Bank of Japan*. The sample period is 1980:1-2004:3 for both series.

B Deconvoluting the Probability Density Function of $\hat{\lambda}$

This appendix describes the procedure we use in section 3.2 to deconvolute the probability density function of $\hat{\lambda}$. To fix ideas, let's start by considering the construction of a (1-α)% confidence interval for $\hat{\lambda}$, $[\hat{\lambda}^L_{(1-\alpha)}, \hat{\lambda}^U_{(1-\alpha)}]$, and let's assume, for the sake of simplicity, that λ_j and $\hat{\lambda}$ can take any value over $[0; \infty)$. Given the duality between hypothesis testing and the construction of confidence intervals, the (1-α)% confidence set for $\hat{\lambda}$ comprises all the values of λ_j that cannot be rejected based on a two-sided test at the α% level. Given that an increase in λ_j automatically shifts the PDF of $\hat{L}_j$ conditional on λ_j upwards, $\hat{\lambda}^L_{(1-\alpha)}$ and $\hat{\lambda}^U_{(1-\alpha)}$ are therefore such that

$$P\left(\hat{L}_j > \hat{L} \mid \lambda_j = \hat{\lambda}^L_{(1-\alpha)}\right) = \alpha/2 \tag{B1}$$

$$P\left(\hat{L}_j < \hat{L} \mid \lambda_j = \hat{\lambda}^U_{(1-\alpha)}\right) = \alpha/2 \tag{B2}$$

Let $\phi_{\hat{\lambda}}(\lambda_j)$ and $\Phi_{\hat{\lambda}}(\lambda_j)$ be the probability density function and, respectively, the cumulative probability density function of $\hat{\lambda}$, defined over the domain of λ_j. The fact that $[\hat{\lambda}^L_{(1-\alpha)}, \hat{\lambda}^U_{(1-\alpha)}]$ is a (1-α)% confidence interval automatically implies that (1-α)% of the probability mass of $\phi_{\hat{\lambda}}(\lambda_j)$ lies between $\hat{\lambda}^L_{(1-\alpha)}$ and $\hat{\lambda}^U_{(1-\alpha)}$. This in turn implies that $\Phi_{\hat{\lambda}}(\hat{\lambda}^L_{(1-\alpha)})$=$\alpha/2$ and $\Phi_{\hat{\lambda}}(\hat{\lambda}^U_{(1-\alpha)})$=1-$\alpha/2$. Given that this holds for any $0<\alpha<1$, we therefore have that

$$\Phi_{\hat{\lambda}}(\lambda_j) = P\left(\hat{L}_j > \hat{L} \mid \lambda_j\right) \tag{B3}$$

In this way, based on the *exp*-Wald test statistic, $\hat{L}$, and on the simulated distributions of the $\hat{L}_j$'s conditional on the λ_j's in Λ, we obtain an estimate of the cumulative probability density function of $\hat{\lambda}$ over the grid Λ, let's call it $\hat{\Phi}_{\hat{\lambda}}(\lambda_j)$. Finally, we fit a logistic function to $\hat{\Phi}_{\hat{\lambda}}(\lambda_j)$ via non-linear least squares and we compute the implied estimate of $\phi_{\hat{\lambda}}(\lambda_j)$—call it $\hat{\phi}_{\hat{\lambda}}(\lambda_j)$—scaling its elements so that they sum to one.

C The Monte Carlo Integration Procedure for Computing Confidence Bands for the Estimated State Vector

This appendix describes the procedure we use in section 3.2 to compute confidence bands for estimated time-varying equilibrium levels of labor productivity growth taking into account of both parameter and filter uncertainty. The procedure is an adaptation to the case at hand of the Monte Carlo integration procedure proposed by Hamilton (1985, 1986).

The first step consists in integrating out parameter uncertainty, i.e. uncertainty pertaining to the true values of λ, σ^2, and of $\Omega \equiv [\sigma_1^2, \sigma_2^2, ..., \sigma_k^2]'$, the vector of the volatilities for the identified sub-periods. Let $\hat{\lambda}$ and $\hat{\phi}_{\hat{\lambda}}(\lambda_j)$ be the median-unbiased estimate of λ and its estimated deconvoluted discretised probability density function, respectively; let $\hat{\sigma}^2_{OLS}$ be the OLS estimate of σ^2; let $\hat{\Omega}_{OLS}$ be the OLS estimate of Ω; and let $\hat{\theta}_{OLS}$ and $\hat{V}(\hat{\theta}_{OLS})$ be the Hansen (1999) grid bootstrap MUB estimate of θ in (1)—where θ is defined as θ=[μ, ϕ_1, ..., ϕ_p]'—and its Newey and West (1987) estimated covariance matrix. We take 10,000 draws from $\hat{\phi}_{\hat{\lambda}}(\lambda_j)$—let's define the i-th draw as $\tilde{\lambda}_i$—and for each of them we do the following.

- If $\tilde{\lambda}_i$>0, we get a draw $\tilde{\sigma}^2$=[(T-p-1)$\hat{\sigma}^2_{OLS}$]/χ^2_{T-p-1}, and we compute the covariance matrix of η_t in (3) as $\tilde{\lambda}_i^2\tilde{\sigma}^2\hat{Q}$. For each of the $\hat{\sigma}_j^2$ in $\hat{\Omega}_{OLS}$, we then get a draw $\tilde{\sigma}_j^2$=[(T_j-p-1)$\hat{\sigma}_j^2$]/$\chi^2_{T_j-p-1}$, thus getting the vector $\tilde{\Omega}$ for the simulated volatilities for each of the identified sub-periods. Conditional on $\tilde{\lambda}_i^2\tilde{\sigma}^2\hat{Q}$ and $\tilde{\Omega}$, we run the Kalman filter and smoother for (2)-(3), thus getting estimates of the state vector and of its precision matrix at each t, $\theta^i_{t|\tau}$ and $P^i_{t|\tau}$, respectively, with τ=t for one sided estimates, and τ=T for two-sided ones.

- If $\tilde{\lambda}_i$=0, we simply set $\theta^i_{t|\tau}$=$\hat{\theta}_{OLS}$ and $P^i_{t|\tau}$=$\hat{V}(\hat{\theta}_{OLS})$ for each t, with τ=t, T.

Finally, for each t we take the mean across the 10,000 draws for both $\theta^i_{t|\tau}$ and $P^i_{t|\tau}$, τ=t, T—let's define them as $\bar{\theta}_{t|\tau}$ and $\bar{P}_{t|\tau}$, respectively—thus integrating out uncertainty about λ and Ω.

The second step then consists in quantifying the extent of filter uncertainty, which we do by repeating the following 10,000 times. For each t from p+1 to T, draw from $MN(\bar{\theta}_{t|\tau}, \bar{P}_{t|\tau})$, τ=t, T, where $MN(h, H)$ is a multivariate normal distribution with mean h and covariance matrix H. Call this draw $\theta^k_{t|\tau}$. Based on $\theta^k_{t|\tau}$, compute the time-varying mean of the series, $\gamma^k_{t|\tau} \equiv \mu^k_{t|\tau}$/(1-$\rho^k_{t|\tau}$). Based on the distribution of the $\gamma^k_{t|\tau}$'s, we then compute both a median estimate of γ (the black lines in Figures 5-7), and 90% confidence bands around the median. Finally, based on a single pass of the Kalman filter and smoother conditional on the MUB estimate of λ and the OLS

estimate of Ω, we compute the 'traditional' estimate of γ found in most applications of the Stock-Watson methodology—see, e.g. Roberts (2001)—which abstracts from parameter uncertainty.

D Computing Confidence Intervals for Cochrane's Variance Ratio Estimator *via* Spectral Bootstrapping

This appendix describes the spectral bootstrapping procedure we use in section 6 to compute confidence intervals for Cochrane's (1988) variance ratio estimator. Let $\tilde{x}_t$ be the discrete Fourier transform of Δy_t, i.e. $\tilde{x}_t$=$a(\omega_j)$-$ib(\omega_j)$, where i is the imaginary number, the ω_j's are the Fourier frequencies, and $a(\omega_j)$ and $b(\omega_j)$ are the Fourier coefficients corresponding to the Fourier frequency ω_j. As it is well known (see, e.g., Brillinger (1981))

$$\frac{a(\omega_j)}{\sqrt{f(\omega_j)}}, \frac{b(\omega_j)}{\sqrt{f(\omega_j)}} \overset{asy}{\rightarrow} iid\ N(0,1/2) \tag{D1}$$

where $f(\omega_j)$ is the spectral density of Δy_t. Following Berkowitz and Kilian (2000), we generate pseudo-Fourier coefficients according to

$$a^*(\omega_j) = \sqrt{\hat{f}(\omega_j)} z_a(\omega_j) \quad b^*(\omega_j) = \sqrt{\hat{f}(\omega_j)} z_b(\omega_j) \tag{D2}$$

where $\hat{f}(\omega_j)$ is a consistent, i.e., smoothed, estimator of the spectral density of Δy_t,[33] and $z_a(\omega_j)$ and $z_b(\omega_j)$ are $iid\ N(0,1/2)$. We then inverse-Fourier transform $\tilde{x}_t^*$=$a^*(\omega_j)$-$ib^*(\omega_j)$, thus getting artificial, boostrapped $\tilde{x}_t^*$'s, and based on them we compute boostrapped $\hat{V}_k^*$, thus building up the empirical distribution of $\hat{V}_k$. In what follows we use 10,000 bootstrap replications. Finally, we compute the α% confidence bands based on the $\alpha/2$ and (1-α)/2 quantiles of the empirical distribution of the $\hat{V}_k^*$.

In order to gauge an idea of the coverage properties of the proposed spectral bootstrap procedure, we perform a simple Monte Carlo experiment based on the ARIMA(0,1,1) process $\Delta y_t = u_t + \theta u_{t-1}$. It can be easily shown that for such a process the variance ratio at horizon k is equal to:

$$V_k = 1 + 2k^{-1}(k-1)\theta(1+\theta^2)^{-1} \tag{D3}$$

which, for $k \rightarrow \infty$, converges to V_∞=$1+2\theta(1+\theta^2)^{-1}$, equal to the fraction of the variance of Δy_t due to the innovation in the permanent component within the Beveridge-Nelson decomposition. We consider values of θ such as to give rise to three values

[33] We estimate $\hat{f}(\omega_j)$ by smoothing the periodogram in the frequency domain by means of a Bartlett spectral window. Following Berkowitz and Diebold (1998), we select the bandwidth automatically via the procedure introduced by Beltrao and Bloomfield (1987).

of V_∞, 0.2, 1—corresponding to the case of a pure random walk—and 1.5. For any of them we first derive the distribution of Cochrane's variance ratio estimator—expression (6) in the text—at horizons up to 80 quarters, based on 10,000 replications of the process, for simulated samples of length 160 quarters. The first row in figure 5 shows, for any of the three values of V_∞, the median and the mean of the simulated distribution of $\hat{V}_k$, the upper and lower 90% percentiles, and the theoretical value of V_k based on (C3). We then simulate the process 1,000 times, and based on each simulation we compute the upper and lower 90% confidence bands for the estimate of V_k based on the previously described spectral bootstrapping procedure (we set the number of bootstrapping replications to 1,000), thus building up their empirical distributions. The second row of figure 5 reports the upper and lower 90% percentiles of the simulated distribution of $\hat{V}_k$—the same shown in the corresponding panels in the first row—together with the means of the distributions of the bootstrapped upper and lower 90% confidence bands. A comparison between the simulated percentiles of the distribution of V_k and the means of the distributions of the bootstrapped confidence bands—i.e., confidence bands's expected values—allows us to get an idea of the accuracy of the proposed procedure. As the three panels in the second row show, the accuracy of the approximation is quite good not only at the 20-year horizon, but also at shorter horizons, with the partial exception of the V_∞=0.2 case.

Table 1 Tests for multiple breaks at unknown points in the sample in the mean based on Andrews and Ploberger (1994) and Bai (1997a)

Break dates and 90% confidence intervals	*exp*-Wald (bootstrapped *p*-value)[a]	Sub-periods	Mean (standard error)
(a) Output per hour			
United States, output per hour, business sector:			
1966:2 [1966:25; 1966:3]	2.37 (0.057)	1947:2-1966:1	3.48 (0.42)
		1966:2-2005:4	2.06 (0.25)
United States, output per hour, manufacturing sector, overall:			
2001:4 [2001:3; 2002:1]	7.16 (0.038)	1987:2-2001:3	3.23 (0.36)
		2001:4-2005:4	5.77 (0.46)
Eurozone, real GDP per hour:			
1995:2 [1995:1; 1995:3]	9.00 (0.008)	1980:2-1995:1	2.33 (0.18)
		1995:2-2004:4	0.93 (0.23)
(b) Output per worker			
Eurozone, real GDP per worker:			
1977:1 [1976:4; 1977:3]	8.21 (0.011)	1970:2-1976:4	3.61 (0.62)
1998:1 [1997:4; 1998:2]	10.94 (0.009)	1977:1-1997:4	1.80 (0.18)
		1998:1-2004:4	0.59 (0.24)

[a] *p*-values have been bootstrapped as in Diebold and Chen (1996).
All other series, no identified break date.

Table 2 Tests for multiple breaks at unknown points in the sample in the mean based on Bai and Perron (1998): double maximum tests, and bootstrapped *p*-values[a]

	UDmax	*WDmax*
United States, output per hour:		
nonfarm business sector	6.81 (0.161)	7.64 (0.153)
business sector	8.02 (0.073)	8.99 (0.059)
nonfinancial corporations	12.54 (0.224)	12.54 (0.299)
manufacturing	13.65 (0.064)	13.65 (0.091)
manufacturing, durables	14.03 (0.242)	14.03 (0.306)
manufacturing, non durables	2.99 (0.642)	3.56 (0.630)
United States, output per worker:		
nonfarm business sector	4.68 (0.618)	5.25 (0.662)
business sector	7.34 (0.183)	7.34 (0.258)
nonfinancial corporations	10.50 (0.369)	11.77 (0.395)
manufacturing	7.05 (0.708)	7.53 (0.755)
manufacturing, durables	10.64 (0.424)	10.64 (0.519)
manufacturing, non durables	1.81 (0.933)	2.45 (0.91)
Eurozone, real GDP:		
per hour	13.12 (0.001)	13.12 (0.002)
per worker	16.81 (0.02)	18.84 (0.013)
Japan, real GDP:		
per hour	3.84 (0.096)	3.84 (0.148)
per worker	6.98 (0.135)	7.82 (0.135)
Australia:		
output per hour, market sector	1.93 (0.889)	3.08 (0.754)
output per hour, all industries	2.47 (0.452)	3.95 (0.195)
real GDP per worker	4.31 (0.678)	4.83 (0.73)

[a] *p*-values have been bootstrapped as in Diebold and Chen (1996). Asymptotic 10% critical values are 7.46 for the *UDmax* and 8.20 for the *WDmax* test statistics.

Table 3 Tests for multiple breaks at unknown points in the sample in the mean based on Bai and Perron (1998): *sup*-F(ℓ+1|ℓ) test statistics, and bootstrapped *p*-values[a]

	$F(2\|1)$	$F(3\|2)$	$F(4\|3)$	Estimated break dates
United States, output per hour:				
business sector	7.17 (0.055)	2.12 (0.339)	0.30 (0.856)	1973:1; 1997:1
manufacturing sector, overall	4.26 (0.465)	0.64 (0.944)	0.37 (0.841)	2001:3
Eurozone, real GDP:				
per hour	3.08 (0.355)	1.41 (0.498)	1.10 (0.321)	1995:1
per worker	12.67 (0.026)	1.56 (0.706)	5.27 (0.028)	1976:4, 1997:4

[a] *p*-values have been bootstrapped as in Diebold and Chen (1996). Asymptotic 10% critical values are 8.51, 9.41, and 10.04.

Table 4 Tests for multiple breaks at unknown points in the sample in the mean based on Bai and Perron (1998): estimated mean productivity growth by sub-sample

United States, output per hour:			
business sector	1947:2-1972:4	1973:1-1996:4	1997:1-2005:4
	3.29 (0.35)	1.65 (0.31)	3.06 (0.27)
manufacturing sector, overall	1987:2-2001:2	2001:3-205:4	
	3.28 (0.37)	5.79 (0.49)	
Eurozone, real GDP:			
per hour	1980:2-1994:4	1995:1-2004:4	
	2.34 (0.19)	0.95 (0.23)	
per worker	1970:2-1976:3	1976:4-1997:3	1997:4-2004:4
	3.49 (0.62)	1.85 (0.19)	0.67 (0.22)

Newey and West (1987) standard errors in parentheses.

Table 5 Results based on the Stock-Watson TVP-MUB methodology: *exp*- and *sup*-Wald test statistics, simulated *p*-values, and median-unbiased estimates of λ

	exp-Wald (p-value)	$\hat{\lambda}$	*sup*-Wald (p-value)	$\hat{\lambda}$
United States, output per hour:				
nonfarm business sector	4.16 (0.023)	0.03103	16.84 (0.006)	0.03793
business sector	2.52 (0.095)	0.01897	10.00 (0.085)	0.02069
nonfinancial corporations	2.96 (0.088)	0.02931	13.14 (0.045)	0.04138
manufacturing sector	8.53 (0.001)	0.05172	23.20 (0.001)	0.05172
manufacturing sector, durables	5.97 (0.008)	0.05172	16.55 (0.011)	0.05172
manufacturing sector, non durables	1.89 (0.211)	0.04828	8.18 (0.158)	0.05172
United States, output per worker:				
nonfarm business sector	0.75 (0.708)	0	3.58 (0.788)	0
business sector	1.26 (0.404)	0.00690	6.14 (0.384)	0.00862
nonfinancial corporations	1.36 (0.397)	0.00862	6.73 (0.369)	0.01034
manufacturing sector	4.69 (0.023)	0.05172	16.81 (0.009)	0.05172
manufacturing sector, durables	4.69 (0.023)	0.05172	16.77 (0.011)	0.05172
manufacturing sector, non durables	5.27 (0.015)	0.05172	6.90 (0.255)	0.04138
Eurozone, real GDP:				
per hour	14.048 (0)	0.05172	32.92 (0.001)	0.05172
per worker	9.15 (0.001)	0.05172	24.85 (0)	0.05172
Japan, real GDP:				
per hour	1.12 (0.454)	0.01379	4.87 (0.500)	0.00345
per worker	1.19 (0.442)	0.01552	6.27 (0.326)	0.02586
Australia:				
output per hour, market sector	2.08 (0.165)	0.03448	7.40 (0.210)	0.03103
output per hour, all industries	16.84 (0)	0.05172	42.26 (0.000)	0.05172
real GDP per worker	0.73 (0.693)	0	4.48 (0.605)	0

Table 6 Tests for multiple breaks at unknown points in the sample in the innovation variance based on Andrews-Ploberger[a] (1994) and Bai (1997a)

Break dates and 90% confidence intervals	*exp*-Wald (p-value)	Sub-periods	Variance,[b] and 90% confidence interval
(a) Output per hour			
United States, output per hour, nonfarm business sector:			
1983:3 [1973:3; 1993:3]	7.79 (0)	1947:2-1983:2	1.02 [8.8E-1; 1.19]
		1983:3-2005:4	0.38 [3.3E-1; 0.45]
United States, output per hour, business sector:			
1982:2 [1973:2; 1991:2]	9.82 (0.001)	1947:2-1982:1	1.12 [9.7E-1; 1.31]
		1982:2-2005:4	0.38 [3.3E-1; 0.45]
(b) Output per worker			
United States, output per worker, nonfarm business sector:			
1983:4 [1975:4; 1991:4]	11.23 (0)	1947:2-1983:3	1.30 [1.12; 1.53]
		1983:4-2003:4	0.40 [3.42; 0.47]
United States, output per worker, business sector:			
1982:2 [1974:1; 1990:3]	11.92 (0)	1947:2-1982:1	1.33 [1.15; 1.57]
		1982:2-2003:4	0.41 [3.5E-1; 0.48]
Eurozone, real GDP per hour:			
1992:3 [1986:1; 1999:1]	10.27 (0)	1970:2-1992:2	0.40 [3.4E-1; 0.47]
		1992:3-2004:4	0.09 [7.7E-2; 0.10]

[a] Identical results based on the Andrews (1993) *sup*-Wald statistic are available upon request. All other series, no identified break date.

[b] Innovation variance is for the rate of growth of labor productivity computed as the quarter-on-quarter rate of change of the relevant index.

Table 7 Power of the tests for breaks in the mean conditional on taking the estimated Stock-Watson TVP-MUB models as data generation processes[a]

	Based on:	
	Andrews and Ploberger's (1994) *exp*-Wald test statistic[b]	Bai and Perron's (1998) *WDmax* test statistic[b]
United States, output per hour:		
nonfarm business sector	0.329	0.390
nonfinancial corporations	0.319	0.310
Australia, output per hour, all industries	0.374	0.434

[a] Critical values have been bootstrapped as in Diebold and Chen (1996).
[b] The test is for a single break at an unknown point in the sample.

Table 8 Cochrane's (1988) variance ratio estimator at the 20-year horizon, and 90% confidence interval (percentage points)

		Bootstrapped distribution of $\hat{V}^*_{80}$		
	$\hat{V}_{80}$	Median	Mean	90% confidence interval
		Output per hour:		
United States:				
nonfarm business sector	2.82	3.95	4.45	[2.26; 8.30]
business sector	1.87	2.80	3.27	[1.77; 6.35]
nonfinancial corporations	5.19	3.92	4.53	[2.49; 8.65]
		Output per worker:		
United States:				
nonfarm business sector	2.44	3.11	3.60	[2.02; 6.83]
business sector	2.04	3.17	3.65	[2.01; 6.99]
nonfinancial corporations	5.21	4.32	4.91	[2.76; 9.04]

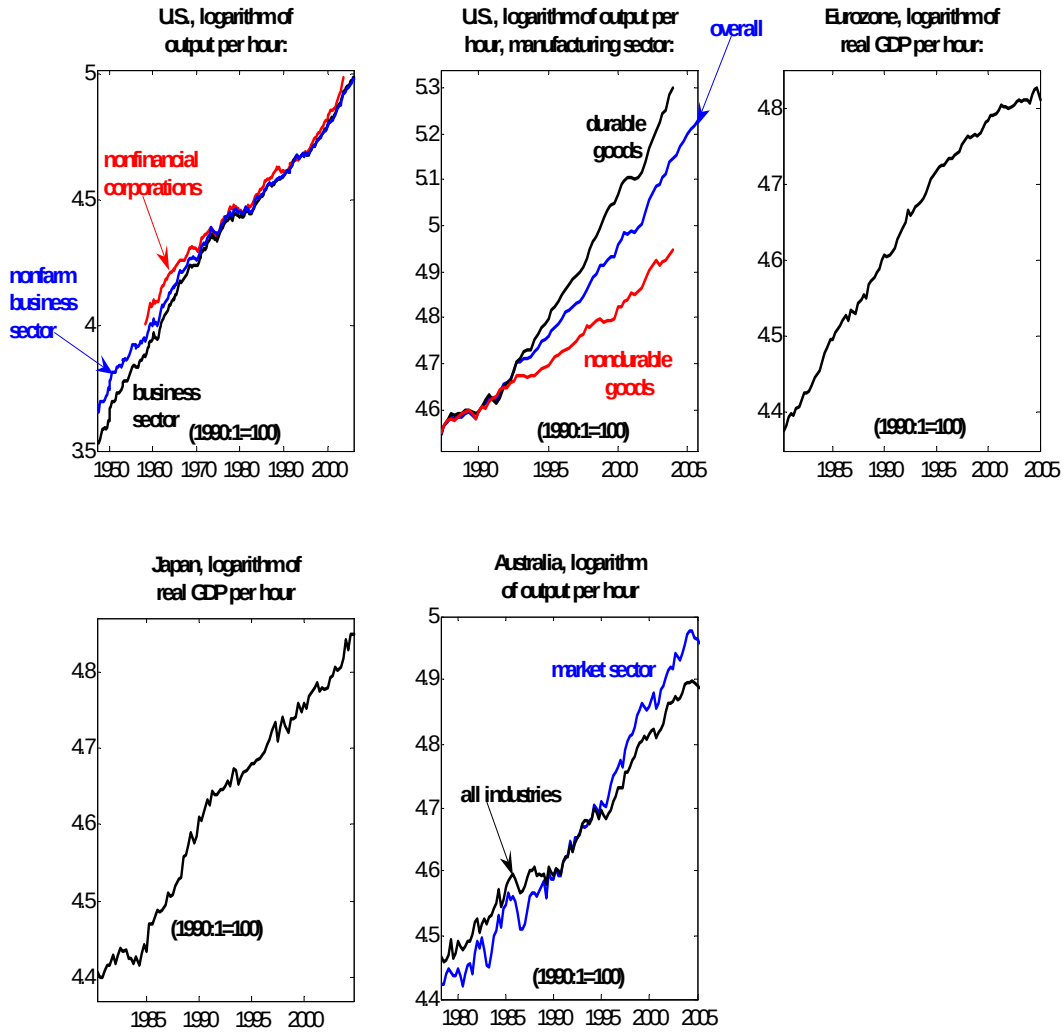

Figure 1: The raw data: logarithm of output per hour in the United States, the Eurozone, Australia, and Japan

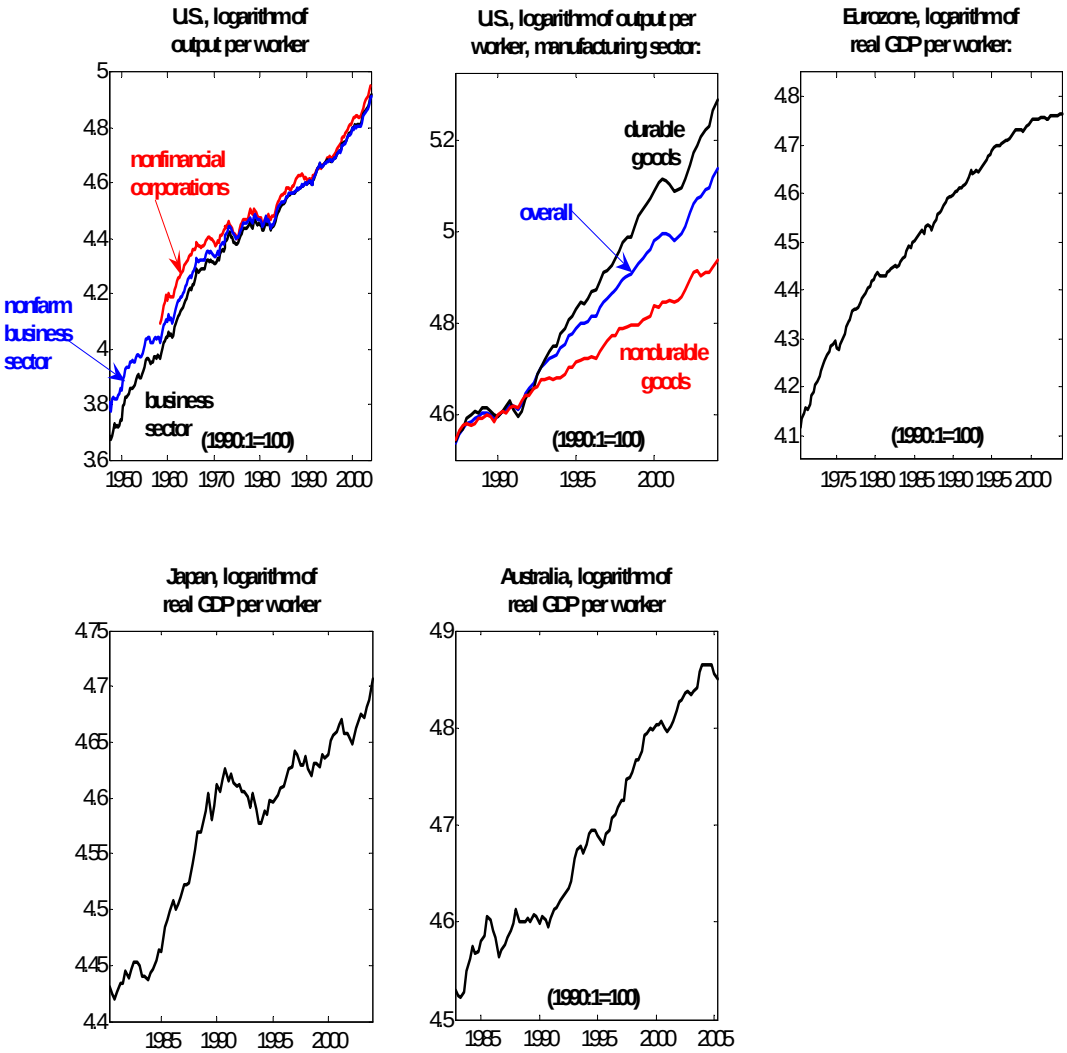

Figure 2: The raw data: logarithm of output per worker in the United States, the Eurozone, Australia, and Japan

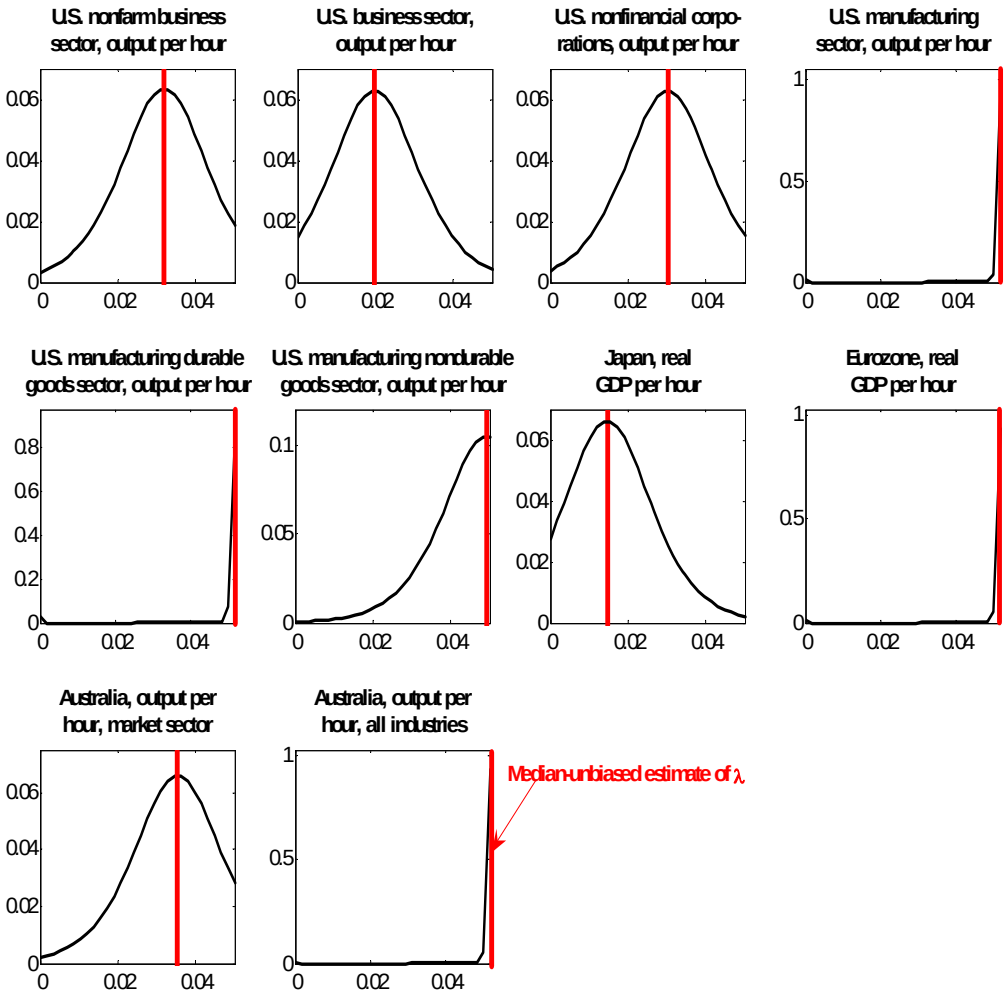

Figure 3: Median-unbiased estimates of λ, and deconvoluted PDFs of $\hat{\lambda}$ (output per hour)

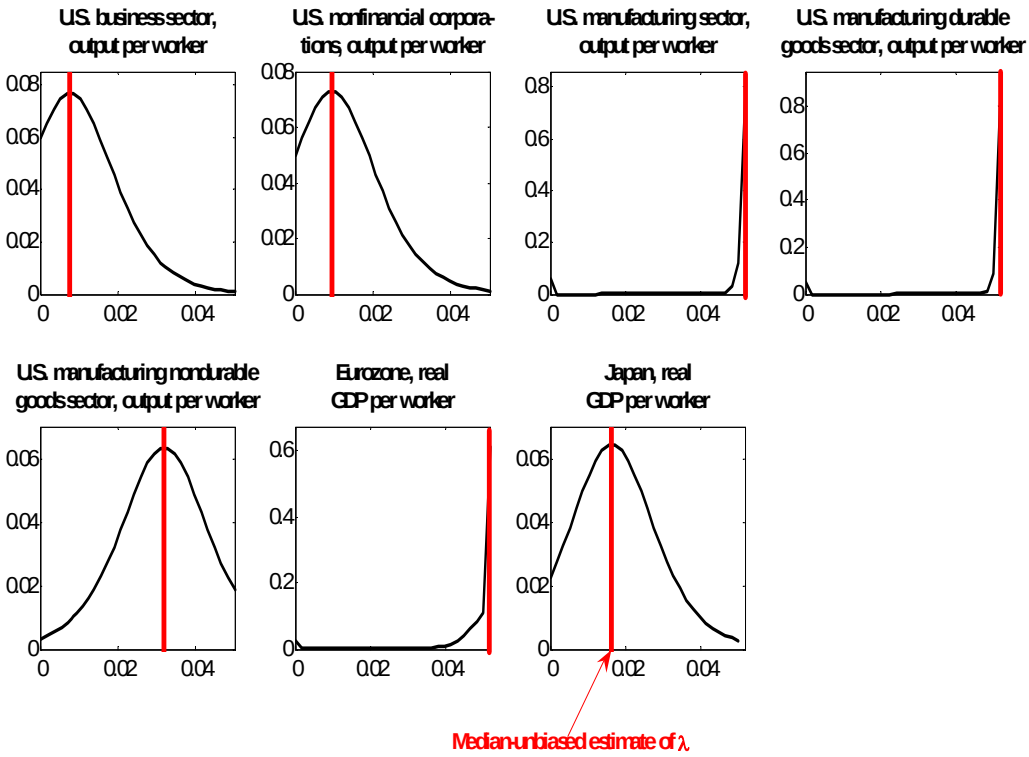

Figure 4: Median-unbiased estimates of λ, and deconvoluted PDFs of $\hat{\lambda}$ (output per worker)

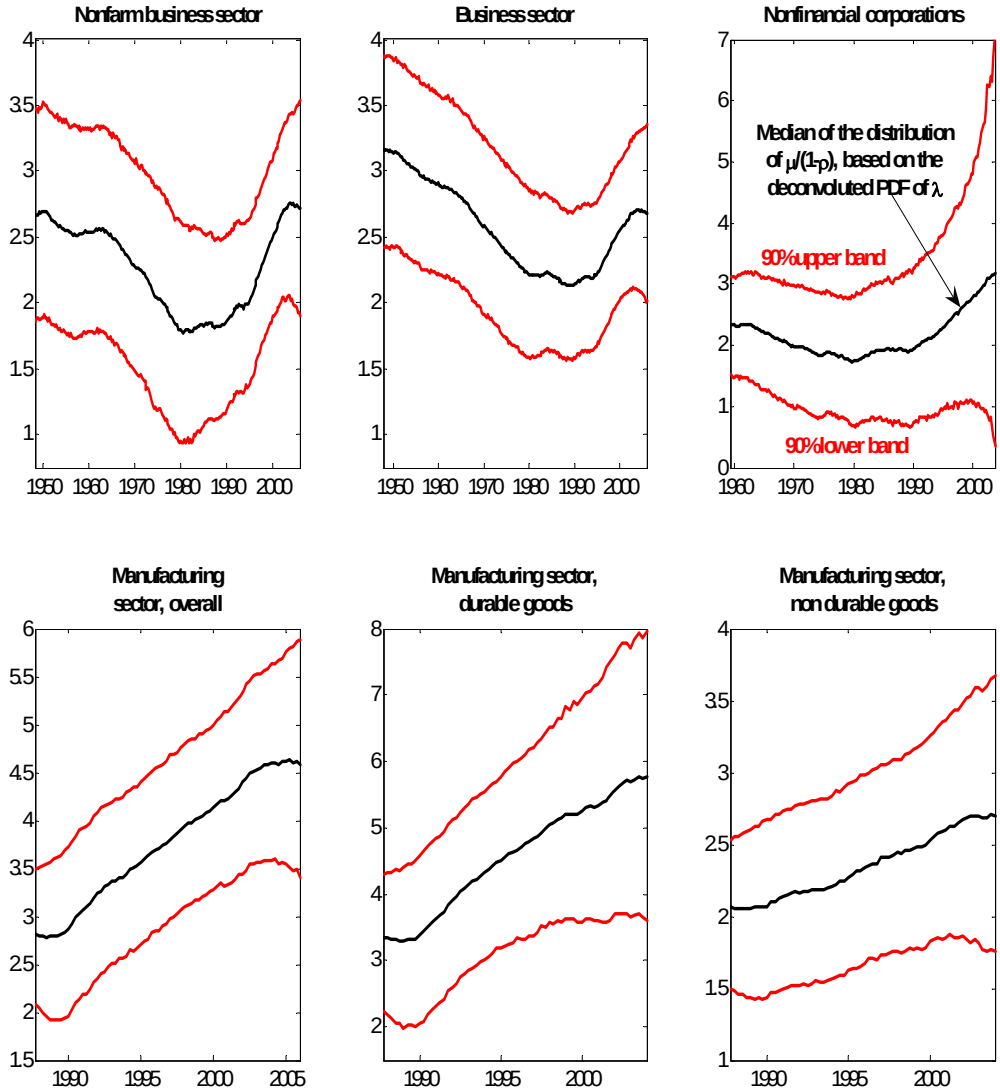

Figure 5: TVP-MUB two-sided estimates of equilibrium U.S. labor productivity growth, and 90% confidence bands (output per hour)

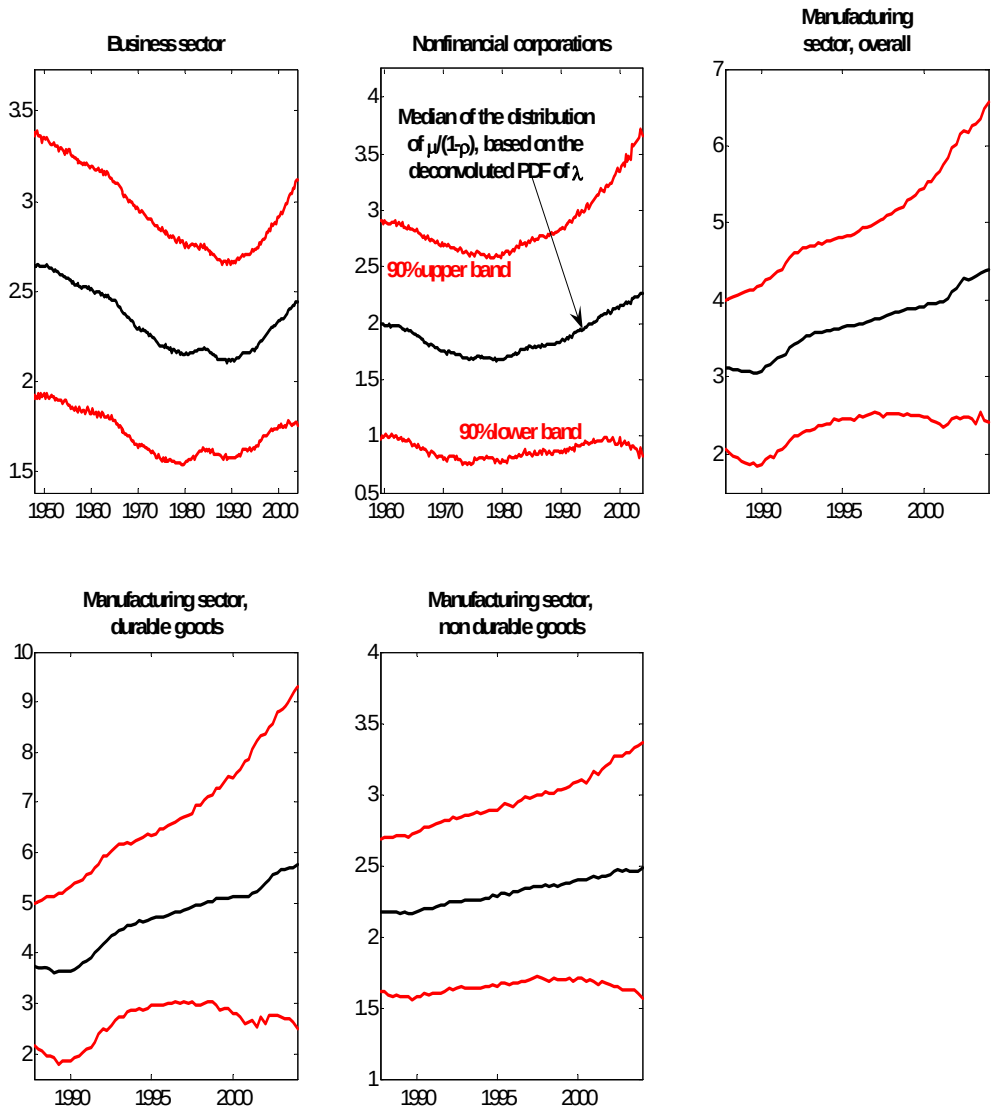

Figure 6: TVP-MUB two-sided estimates of equilibrium U.S. labor productivity growth, and 90% confidence bands (output per worker)

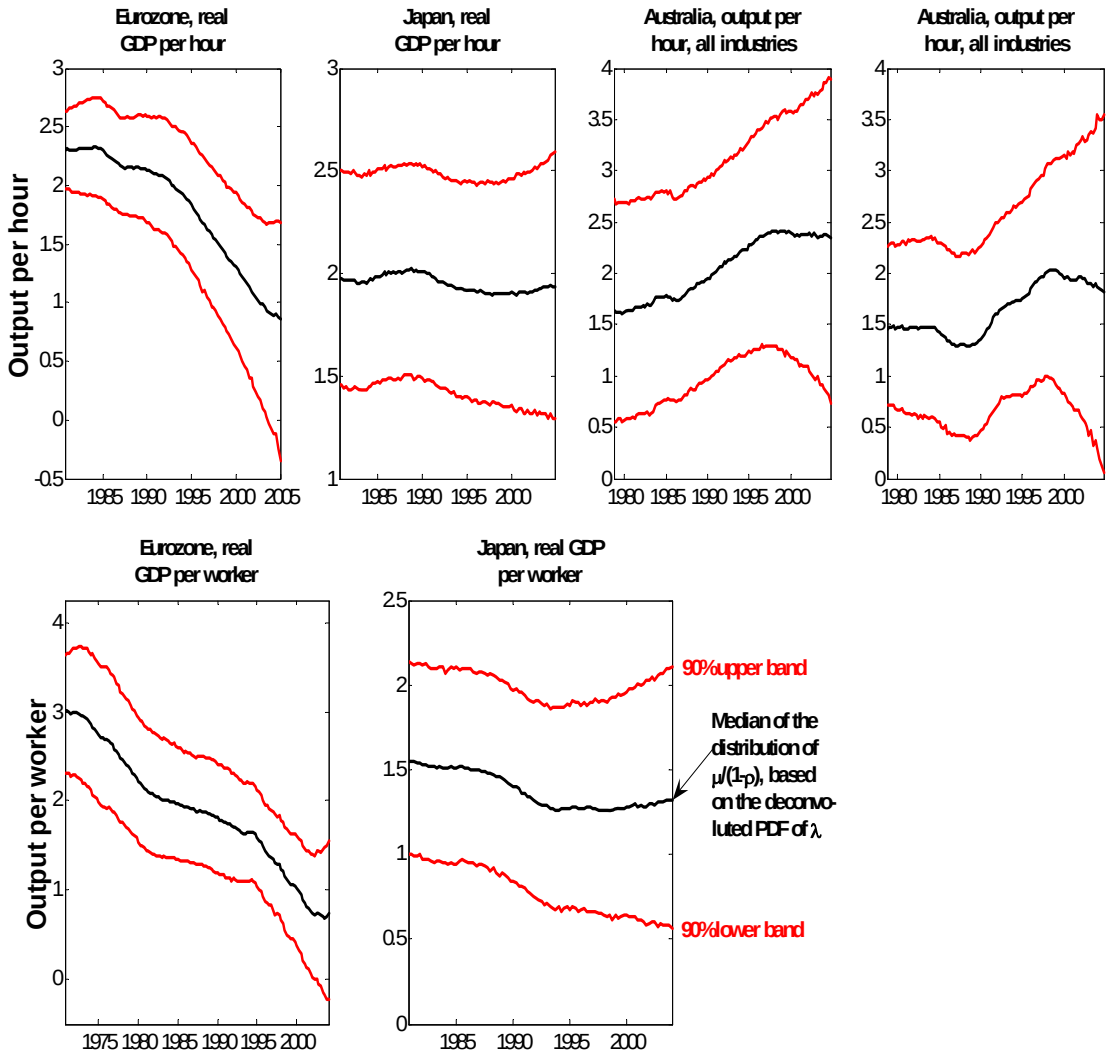

Figure 7: TVP-MUB two-sided estimates of equilibrium labor productivity growth in the Eurozone, Japan, and Australia, and 90% confidence bands (output per hour and output per worker)

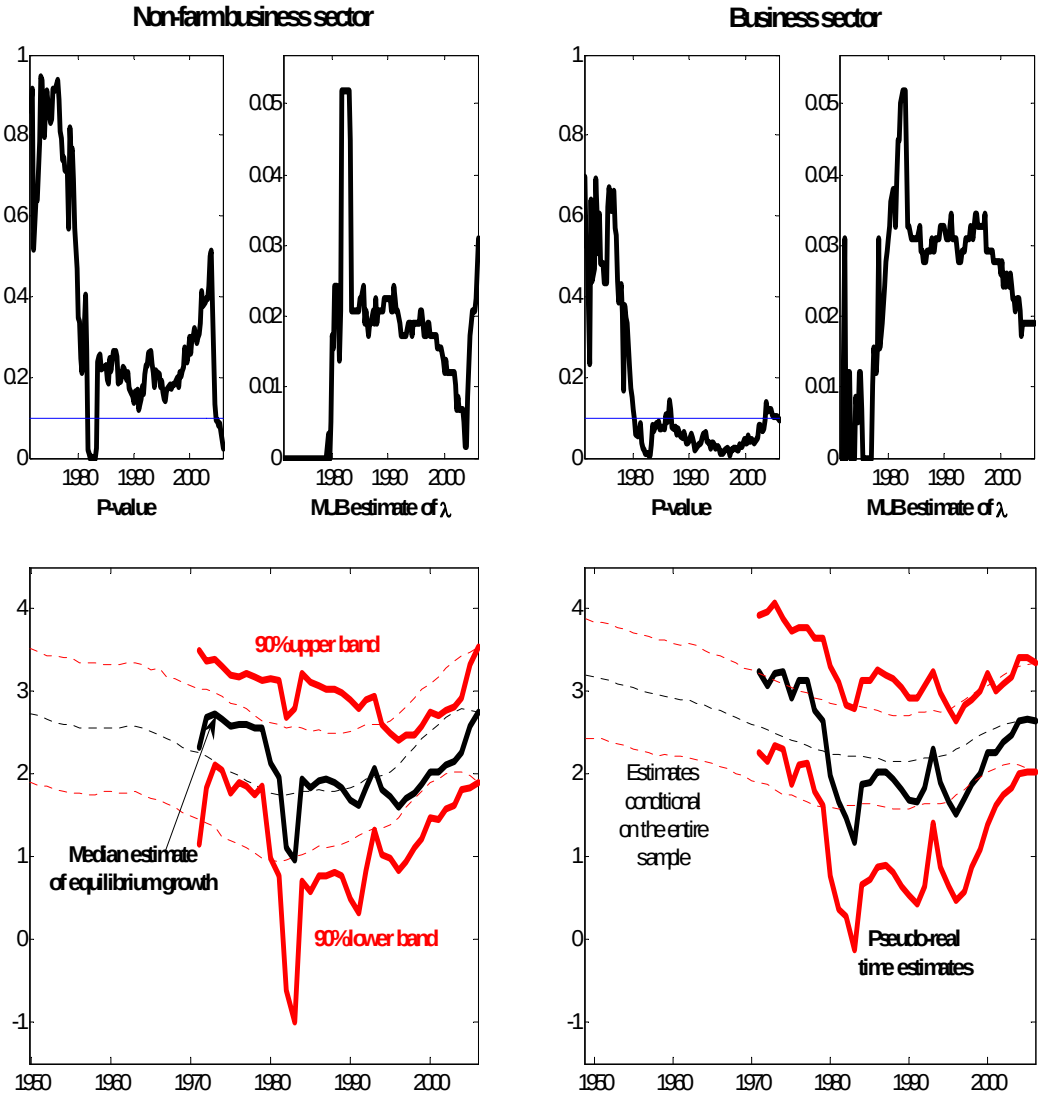

Figure 8: Pseudo-real time estimates of equilibrium productivity growth in the United States

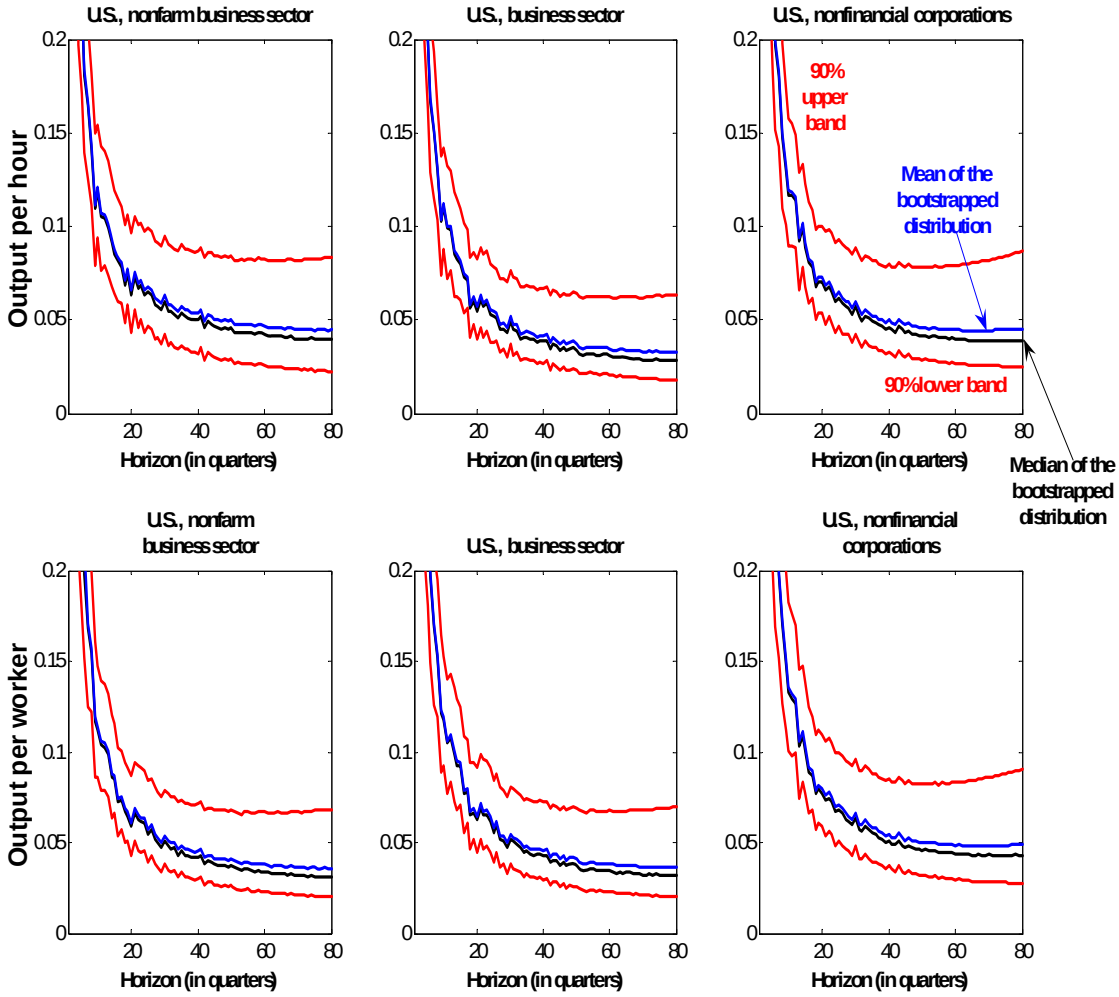

Figure 9: United States, estimates of the permanent component of the quarter-on-quarter change in the rate of growth of labor productivity: results based on Cochrane's variance ratio estimator at the 20-year horizon

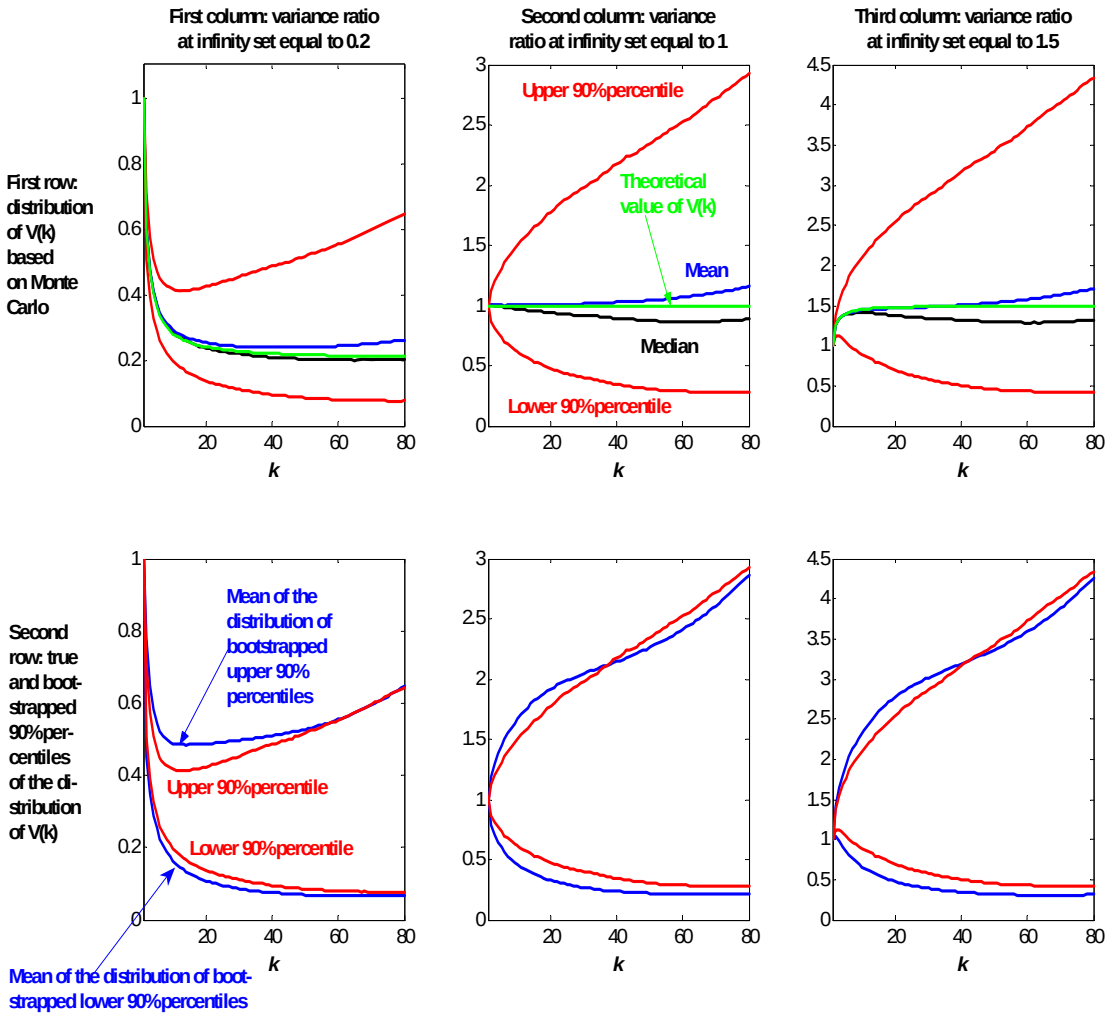

Figure 10: Monte Carlo evidence on the performance of Cochrane's (1988) variance ratio estimator, and on the accuracy of the confidence intervals based on spectral bootstrapping