EUROPEAN CENTRAL BANK
EUROSYSTEM

WORKING PAPER SERIES
NO 718 / JANUARY 2007

DRIFT AND BREAKS IN LABOR PRODUCTIVITY

by Luca Benati

EUROPEAN CENTRAL BANK

EUROSYSTEM

WORKING PAPER SERIES

NO 718 / JANUARY 2007

DRIFT AND BREAKS IN LABOR PRODUCTIVITY [1]

by Luca Benati [2]

In 2007 all ECB publications feature a motif taken from the €20 banknote.

1 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 and the Editor of the Journal of Economic Dynamics and Control (Peter Ireland) 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 European Central Bank. Part of this work was done while I was visiting the Directorate General Research of the European Central Bank, whose kind hospitality is gratefully acknowledged.

2 Monetary Policy Strategy Division, European Central Bank, Kaiserstrasse 29, 60311 Frankfurt am Main, Germany; e-mail: Luca.Benati@ecb.int

Address
Kaiserstrasse 29
60311 Frankfurt am Main, Germany

Postal address
Postfach 16 03 19
60066 Frankfurt am Main, Germany

Telephone
+49 69 1344 0

Internet
http://www.ecb.int

Fax
+49 69 1344 6000

Telex
411 144 ecb d

ISSN 1561-0810 (print)
ISSN 1725-2806 (online)

CONTENTS

Abstract

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, the United Kingdom, 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. 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 equilibrium productivity growth stabilising over the most recent period.

Keywords: Structural break tests; time-varying parameters; median-unbiased estimation; variance ratio; bootstrapping; frequency domain; Monte Carlo integration.

JEL codes: E30, E32.

Non Technical Summary

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. 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, Athanasios Orphanides has argued, for example, that part of the blame for the Great Inflation of the 1970s 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.

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 (TVP-MUB) 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, the United Kingdom, Australia, and Japan over the post-WWII period. Our main results be summarised as follows.

Based on either the Bai and Perron (1998, 2003) methodology, or the Bai (1997) 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 the sometimes surprisingly low power exhibited by state-of-the-art break tests when the true data generation process (DGP) is characterised by random-walk time-variation—a simple way of expressing the notion of 'slow and continuous change'. To put it differently, historical changes in equilibrium productivity growth have most likely been too gradual to be detectable *via* such a powerful, but ultimately quite crude methodology.

The more flexible Stock and Watson's 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, until the first half of the 1960s, with trend productivity growth estimated between 2.4% and 2.6%; a marked slowdown from the first half of the 1960s up to around 1980, with equilibrium productivity growth estimated to have fallen to 1.9% in 1982:1; a period of stagnation, the 1980s, with trend growth fluctuating between 1.9% 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 sam-

ple, 2006:3, at 2.4%. Interestingly, evidence clearly points towards the productivity acceleration to have reached a *plateau* over the last few years, and, if anything, to have experienced a very mild decrease over the last three years. Results for the Euro-zone point towards a marked deceleration since the beginning of the 1980s, with the equilibrium rates of growth of output per hour and per worker stabilising, over the last few years, at 1.1% and 0.8% respectively. As for the other contries, we detect no evidence of time-variation in the United Kingdom, and some mild evidence 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.

Based on a pseudo-real time exercise, we then explore the issue of whether the Stock and Watson's TVP-MUB methodology might allow to detect productivity slowdowns and speedups as they take place. Our evidence 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. By the same token, the methodology would have failed to detect the productivity acceleration of the second half of the 1990s.

We conclude by stressing three key messages a policymaker should take home from this paper. First, equilibrium (trend) labor productivity growth should be regarded, *in general*, as time-varying. Second, given that, as we show, 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. Finally, when time-variation in equilibrium productrivity growth does take place, it takes place *most likely gradually*, i.e. without sudden jumps, so that the best way of analysing it is *via* time-varying parameters models, rather than *via* break tests.

1 Introduction[1]

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[2] 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 the U.S. nonfarm business and business sectors estimated at 2.4% and 2.5%, respectively, in 2006:3—assuming a trend rate of growth of productivity in the business sector of 1.3% appears indeed as unduly pessimistic.[3]

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,[4] 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

[1] This paper updates the estimates contained in Benati, L. (2007), 'Drift and Breaks in Labor Productivity', *Journal of Economic Dynamics and Control*, forthcoming.

[2] See Gordon (1999).

[3] 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. In fact, as we discuss more extensively in section 3.2.1 below, our results clearly suggest that the U.S. productivity resurgence of the second half of the 1990s has reached a *plateau* around the very beginning of the new century, and even point towards the possibility of a *deceleration* taking place as we write.

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

an over-estimation of the authentic amount of slack existing in the economy.[5] 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,[6] so that failure to identify shifts in equilibrium productivity growth automatically leads to a mis-estimation of the natural rate of interest, with potentially dire consequences for monetary policy.[7]

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, the United Kingdom, Australia, and Japan over the post-WWII period. Our main results be summarised as follows.

- 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 the sometimes surprisingly low power exhibited by state-of-the-art break tests when the true data generation process (henceforth, DGP) is characterised by random-walk time-variation—a simple way of expressing the notion of 'slow and continuous change' in the underlying DGP. To put it differently, historical changes in equilibrium productivity growth have most likely 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

[5] 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).

[6] 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.

[7] The recent work of Trehan and Wu (2006) contains however a qualification to Orphanides' position. As Trehan and Wu point out, to the extent that the equilibrium real rate '[...] moves in the same direction as the trend growth rate (as is suggested by theory), the probability that an unperceived change in trend growth will lead to a substantial change in inflation is noticeably lower than is suggested by recent analyses that assume a constant ERR.' The intuition is that '[...] slower trend output growth is accompanied by a lower equilibrium interest rate, so that a policy authority which takes a while to recognize that the growth rate of output has slowed down will also take a while to recognize that the equilibrium interest rate has fallen. Thus, policy will, in fact, not be as stimulative as would be suggested by analyses that ignore the link between trend output growth and the ERR [equilibrium real rate], and inflation will not rise as much.'

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 rates of growth of output per hour and per worker stabilising, over the last few years, at 1.1% and 0.8% respectively. As for the other contries, we detect no evidence of time-variation in the United Kingdom, and some mild evidence 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.

1.1 A comparison with the previous literature

With the notable exceptions of Ben-David and Papell (1998) and Gust and Marquez (2000),[8] the recent literature on productivity growth—see in particular Gordon (1999), Oliner and Sichel (2000), Hansen (2001), Roberts (2001), Oliner and Sichel (2002), Gordon (2003), and Basu, Fernald, Oulton, and Srinivasan (2003)—has exclusively concentrated on the United States,[9] trying to ascertain, first, whether the productivity acceleration which started in the second half of the 1990s is indeed for real, and second, what its likely causes might be. Overall, the broad consensus emerging from this literature is that (1) the 1990s' productivity acceleration is indeed to a very large extent authentic—i.e., it reflects genuine productivity improvements, rather than cyclical factors, or simply an unsustainable pace of capital accumulation—and (2) it appears to be largely attributable to the impact of information technology.

In line with Ben-David and Papell (1998) and Gust and Marquez (2000), on the other hand, the present work explicitly adopts an international dimension, by studying, as we stress in Appendix A, all countries for which we could find reliable data. Our *ex post* dominant focus on the United States therefore uniquely reflects the fact that (*i*) the U.S. possesses the best and most extensive collection of productivity data, and (*ii*) for some countries, in particular the United Kingdom and Japan, either we do not identify any time-variation in trend productivity growth, or we do identify a very modest extent of time-variation. In spite of the overall broad similarity coming from the common international focus, the present work differs however from that of Ben-David and Papell (1998) and Gust and Marquez (2000) along a number of dimensions.

[8] Basu, Fernald, Oulton, and Srinivasan (2003) constitute a partial exception, as they investigate the causes underlying the U.S. productivity acceleration of the second half of the 1990s, and the lack of an analogous acceleration in the United Kingdom.

[9] See also Kouparitsas (2005).

First, strictly speaking, the work of Gust and Marquez (2000) is not based on econometric techniques, as they simply compare average productivity growth rates across countries and sub-periods. Based on (broadly) comparable OECD data, they show how (*i*) during the period between the beginning of the 1980s and mid-1990s, labor productivity growth in the United States had consistently been lower than in all other members of the G-7, but (*ii*) since mid-1990s the opposite has held true, with the United States having consistently outperformed all other members of the G-7.

Ben-David and Papell (1998) search for evidence of changes in the rate of growth of annual real GDP *per capita* for a sample of 74 countries over the period 1950-1990, based on the Summers and Heston (1993) dataset and the battery of break tests developed by Vogelsang (1997). They detect evidence of a growth slowdown—defined as 'statistically significant negative break in the trend function of the growth process'—for 54 countries. Specifically, most advanced countries, with the notable exception of the U.S., the U.K., and Canada, exhibit a slowdown around the time of the first oil shock, while most developing countries appear to have experienced a decrease in the rate of growth of real GDP *per capita* between 1978 and 1983, around the time of the second oil shock and of the beginning of the emerging markets debt crisis. Strikingly, for several of the countries belonging to this second group the slowdown appears to have been an authentic *meltdown*, with the post-break trend function estimated to have actually been *negative*. Crucial differences between Ben-David and Papell (1998) and the present work are that (*i*) their work is exclusively based on endogenous break tests, while the core of the present work is based on the Stock-Watson time-varying parameters median-unbiased estimation methodology, which—as we argue in section 4 below based on Monte Carlo evidence—should in general be regarded as more reliable;[10] and (*ii*) as we mentioned, they only analyse output per capita (rather than output per hour worked, or output per worker), so that, by definition, their work is only partially informative on international productivity developements.

From a methodological point of view, the only two previous studies based on time-varying parameters techniques that we are aware of are those of Roberts (2001) and Gordon (2003) for the United States. The present work should be regarded as largely complementary to these papers as both of them (*i*) uniquely focus on a *single* measure of productivity, as opposed to the *twelve* U.S. series we analyse here; and (*ii*) they are based on semi-structural *multivariate* models—linking changes in cyclical productivity to fluctuations in the cyclical components of other macroeconomic series—while the present work is entirely based on univariate methods. It is worth stressing that, in spite of such methodological difference, Roberts' and Gordon's re-

[10] It is important to stress, however, that, based on our own experience, the application of a 'data-hungry' methodology like Stock and Watson's to samples of annual post-WWII data is essentially unfeasible.

sults are remarkably close to those we obtain for the series they analyse, both in terms of the overall pattern of variation and, often, even from a strictly numerical point of view. For example, for the last quarter of his sample, 2001:1, Roberts (2001, pages 17 and 19) estimates an equilibrium rate of growth for output per hour in the nonfarm business sector equal to 2.5%, very close to the 2.6% we estimate for that quarter.[11]

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;[12] 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.[13] We

[11] An important point to stress is that our estimates, compared with Roberts', are based on almost five more years' of revised data.

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

[13] 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),[14] setting the number of bootstrap replications to 1,000. Finally, we compute confidence intervals for estimated break dates according to Bai (1997b).

Tables 1-2 report the results based on the Andrews-Ploberger *exp*-Wald and the Andrews *sup*-Wald test statistics. Quite surprisingly, we identify break dates[15] for only five series, output per hour in the U.S. business, nonfinancial corporations, and manufacturing durables sectors, and real GDP per hour and per worker in the Eurozone. Results for the three U.S. series for which we detect breaks well accord with the conventional wisdom notion of a 'golden era' of comparatively high productivity growth, the 1950s and 1960s; a significant slowdown beginning in the first half of the 1970s; and a strong growth resurgence starting around mid-1990s. For the business sector, in particular, we identify–broadly in line with Fernald (2005)–two break dates, in 1973:2 and 1995:4, and a U-shaped evolution of mean productivity growth, first falling from 3.3% to 1.4%, and then increasing to 2.9% over the most recent subperiod.[16] Results for the Eurozone point towards a significant productivity slowdown over the sample period, with the equilibrium (mean) rate of growth of real GDP per hour estimated to have fallen from 2.1% before 2001:1 to 0.6% over the most recent sub-period; and the equilibrium growth rate of real GDP per worker having decreased from the 3.3% of the former sub-period to 1.4% over the latter.

Overall, endogenous break tests' failure to detect much evidence of time variation appears as quite puzzling. And especially puzzling is the contrast between the results for the U.S. business sector and those for the nonfarm business sector, traditionally regarded as the 'bellwether series' for U.S. productivity studies.What can account for these results? There are two possible explanations for our failure to identify, overall, much evidence of time-variation, 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

robustness to misspecification in the number of breaks.

[14] To be precise, in order to take into account of possible changes in the volatility of reduced-form innovations over the sample period (we wish to thank an anonymous referee for bringing this point out), we modify the original Diebold and Chen (1996) bootstrap methodology along the following lines. First, we divide the overall sample period into sub-samples based on the results from tests for multiple structural breaks at unknown points in the sample in the innovation variance reported in section 3.2 below. Then, in performing the Diebold and Chen (1996) bootstrap procedure, we simply draw the residuals by sub-period.

[15] The break dates we report correspond to the *beginning* of each sub-sample.

[16] It is important to stress that Fernald's sample period, 1950:2-2004:2, is different from ours, and that he used a different vintage of data.

and without breaks, is more likely to have generated the data, is on the other 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 very close 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 *cum* stochastic volatility 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 3'0% 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),[17] 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 *via* the previously described modification of the Diebold and Chen (1996) procedure, 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 3-5 report the results. As the tables clearly show, compared with the results we saw in the previous paragraph we detect even less evidence of time-variation, with breaks being identified for only three series, output per hour in the U.S. business sector, for which we identify the same break dates we identified in the previous subsection, and output for hour and per worker in the Eurozone, for both of which we identify two breaks, and a monotonic decrease over the sample period. As for the other series, 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.

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

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[18]

$$y_t = \mu + \phi_1 y_{t-1} + \phi_2 y_{t-2} + ... + \phi_p y_{t-p} + u_t = \theta' z_t + u_t \quad (1)$$

where y_t is the rate of growth of labor productivity, θ=[μ, ϕ_1, ..., ϕ_p]$'$ and z_t=[1, $y_{t-1,t}$, ..., $y_{t-p,t}$]$'$. We select the lag order, p, based on the Akaike information criterion,[19] 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, $\phi_{1,t}$, ..., $\phi_{p,t}$]$'$, the time-varying parameters version of (1) is given by:

$$y_t = \theta_t' z_t + u_t \quad (2)$$

$$\theta_t = \theta_{t-1} + \eta_t \quad (3)$$

[18] 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.

[19] An alternative, qualitatively similar set of results based on the Schwartz Information Criterion is not reported here, but is available from the author upon request.

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 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.[20]

3.1 Searching for random-walk time-variation

Our point of departure is the OLS estimate of θ in (1). Conditional on $\hat{\theta}_{OLS}$ we compute the residuals, $\hat{u}_t$, and the estimate of the innovation variance, $\hat{\sigma}^2$, and we perform an *exp*- and a *sup*-Wald joint test for a single break at an unknown point the sample in μ and ρ—with ρ defined as the sum of the AR coefficients in (1)—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 $\lambda_j=0.05$, 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 $\lambda_j=0$.

Table 6 reports the results. Starting from the United States we detect, based on the simulated p-values, strong evidence of time-variation at the 10% level for all the series for output per hour, with the single exception of nonfarm business sector, for which evidence is very strong based on the *sup*-Wald statistic, and is instead not significant based on the *exp*-Wald test. As for output per worker evidence is again

[20] 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.

very strong for nonfinancial corporations and for all the manufacturing series, which it is mixed for the business sector, and not significant for the nonfarm business sector. Turning to other countries, for the United Kingdom we estimate, based on either statistic, λ to be exactly zero for both output per hour and output per worker; 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)[21] 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 1 and 2 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.[22]

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),[23] estimating time-varying parameters models without controlling for the possible presence of heteroskedasticity causes a systematic overestimation of the authentic extent

[21] See also Hamilton (1985).

[22] 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 sector. Although the p-values reported in Table 6 are above 10%, as Figure 3 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.

[23] See also Cogley (2005).

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),[24] 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,[25] 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 7:[26] 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 3-5 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 'bellwether 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 $\hat{\lambda}$—[27]estimated between 2.4% and 2.6%;

[24] 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.)

[25] 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.

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

[27] The reason why this is our preferred estimate is because—different from the one conditional on the TVP-MUB estimates of λ—it takes into account of all possible sources of uncertainty. Estimates

- a marked slowdown from the first half of the 1960s up to around 1980, with equilibrium productivity growth estimated to have fallen to 1.9% in 1982:1;

- a period of stagnation, the 1980s, with trend growth fluctuating between 1.9% 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, 2006:3, at 2.4%. Interestingly, evidence clearly points towards the productivity acceleration to have reached a *plateau* over the last few years, and, if anything, to have experienced a very mild decrease over the last three years.[28]

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, and of a possible deceleration over the last three years, is again very clear—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 2006:3, 2.5%.

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 much stronger acceleration since the beginning of the 1990s, with trend growth increasing from 2.0% in 1990:1 to 3.0% in the last quarter of the sample, 2005:3; and (*ii*) 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 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 2.9% in 1989:4 to 4.1% in 1998:1, and clear evidence of a *plateau* reached ever since; and

- significant differences between the durables and nondurables goods sectors, with trend growth for nondurables increasing from 2.0% at the beginning of the

conditional on the TVP-MUB estimates of λ are however, in general, very close to the median estimates, and are available upon request.

[28] The possibility that U.S. productivity growth has reached a *plateau*, or even that it is currently decelerating, has been recently discussed in the financial press. See e.g. two articles on the October 28, 2006, issue of *The Economist*, 'A Falling Speed Limit' (page 16), and 'Slow Road Ahead' (pages 97-98).

sample to 2.8-2.9% over the most recent years, and with the corresponding figures for the production of durables being 3.1% and 6%, respectively.

Turning to output per worker (Figure 6), 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 nonfarm business sector, for example, trend growth is estimated to have been equal to 2.2% at the beginning of the sample, to have decreased to a minimum of 1.8% in 1980:1, and to have increased to 2.2% in the last quarter of the sample, 2005:4.

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 1.4 at the beginning of the sample, it decreased to a minimum of 1.0 towards the end of the 1990s, and it increased, again, to 1.4 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' conventional wisdom, results based on either real GDP per hour or real GDP per worker point towards a drastic deceleration since the beginning of the 1980s and of the 1970s, respectively. Equilibrium growth of output per hour is estimated to have fallen from 2.4% in 1980:3 to a minimum of 1.1% in 2005:4, with the corresponding figures for output per worker being 3.0% in 1970:3, and 0.8%, in 2006:2. Interestingly, for both series evidence clearly suggests equilibrium productivity growth to have stabilised over the last few years.

Consistent with the comparatively large p-values reported in Table 6, 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 output per hour within all industries has been fluctuating between a minimum of 1.4% and a maximum of 1.9%, with trend 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.[29] In this section we provide some tentative evidence on the plausibility of this conjecture *via* the following Monte Carlo experiment.

We consider two 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 series are output per hour in the U.S. nonfarm business and nonfinancial corporations[30] 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,[31] drawing the pseudo innovations from pseudo random *iid* $N(0, \hat{\sigma}^2)$, and for each simulation (i) we perform 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 8 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.[32] 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, less than 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

[29] 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.

[30] For the U.S. nonfinancial corporations sector the result reported in Table 6 based on the *exp*-Wald test statistic is exactly borderline, with a p-value of 0.10. We have decided to consider this series nonetheless given that the p-value based on the *sup*-Wald test statistic is quite low, being equal to 0.041.

[31] 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.

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

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 were magically catapulted into the Federal Reserve Board sometimes during the first half of the 1970s.[33] 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 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);[34] and (2) the median estimates, together with the 90% confidence bands, conditional on the full samples up to 2006:3, 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

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

[34] 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.

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',[35] 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. Further, when the TVP-MUB methodology finally started 'catching up with reality', it went from a significant *over*estimation to a marked *under*estimation.

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 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. While 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. Further, as a mirror image, once again, of the 1970s, when it finally catched up it went from underestimating equilibrium productivity growth to overestimating it. 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.

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

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

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

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, and on real GDP per worker in the United Kingdom.

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 9 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 9 clearly show how, in both countries, a non-negligible fraction of the quarter-onquarter 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 [2.4; 8.2] 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 home from this paper? Essentially three, we believe.

- First, equilibrium (trend) labor productivity growth should be regarded, *in general*, as time-varying.
- This automatically leads us to a second crucial point. Given that—as we have shown—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. As it is well known, it is part of Chairman Greenspan's legend that he somehow 'picked up' the U.S. productivity acceleration of the second half of the 1990s not *via* 'hard' econometrics, but rather based on his subtle reading of a myriad of economic indicators ...
- Finally, when time-variation in equilibrium productrivity growth does take place, it takes place *most likely 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.

Basu, S., J. Fernald, N. Oulton, and S. Srinivasan (2003): "The Case of the Missing Productivity Growth, Or Does Information Technology Explain Why Productivity Accelerated in the United States But Not in the United Kingdom?," *in M. Gertler and K. Rogoff, eds. (2004), NBER Macroeconomics Annuals 2003.*

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

Ben-David, D., and D. Papell (1998): "Slowdowns and Meltdowns: Postwar Growth Evidence From 74 Countries," *Review of Economic and Statistics*, 80(4), 561–571.

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., 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.

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

——— (2003): "Exploding Productivity Growth: Context, Causes, and Implications," *Brookings Papers on Economic Activity*, 2003:2, 207–279.

Gust, C., and J. Marquez (2000): "Productivity Developements Abroad," *Federal Reserve Bulletin*, 80(4), 665–681.

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.

HANSEN, B. (2001): "The New Econometrics of Structural Change: Dating Changes in U.S. Labor Productivity," *Journal of Economic Perspectives*, 15, 117–128.

JACQUIER, E., N. G. POLSON, AND P. ROSSI (2004): "Bayesian Analysis of Stochastic Volatility Models„" *Journal of Business and Economic Statistic*, 12, 371–418.

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.

KOUPARITSAS, M. A. (2005): "Is There Evidence of the New Economy in U.S. GDP Data?," *Economic Perspectives, Federal Reserve Bank of Chicago*, 1, 12–29.

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.

OLINER, S., AND D. SICHEL (2000): "The Resurgence of Growth in the Late 1990s: Is Information Technology the Story?," *Journal of Econonomic Perspectives*, 14(4), 3–22.

——— (2002): "Information Technology and Productivity: Where Are We Now and Where Are We Going?," *Federal Reserve Bank of Atlanta Economic Review*, Third Quarter, 15–44.

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 and 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.

Summers, R., and A. Heston (1993): "The Penn World Table (Mark 5.5)," .

Trehan, B., and T. Wu (2006): "Time Varying Equilibrium Real Rates and Monetary Policy Analysis," *Journal of Economic Dynamics and Control*, forthcoming.

Vogelsang, T. J. (1997): "Wald-Type Tests for Detecting Shifts in the Trend Function of a Dynamic Time Series," *Econometric Theory*, 13, 818–849.

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 1987:1-2006:3 for the manufacturing sector, and 1947:1-2006:3 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 1958:1-2005:3 for nonfinancial corporations, and 1987:1-2005:4 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 1947:1-2005:4 for the nonfarm business and business sectors; 1958:1-2005:3 for the nonfinancial corporations sector; and 1987:1-2005:4 for the three other series.

A.2 United Kingdom

Quarterly seasonally adjusted series for real GDP per hour and per worker have been computed as the ratios between the quarterly real GDP series (acronym: ABMI) from the *Office for National Statistics* (henceforth, *ONS*), available for the period 1955:1-2006:2, and series from the *Labour Force Survey* for total actual weekly hours worked (YBUS), available since 1971:1, and for employment in the whole economy (TXEL), available since 1959:3, respectively.

A.3 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 Ramon Gomez-Salvador.[36] 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 1970:1-2006:2.

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

A.4 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.

A.5 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 $\theta=[\mu, \phi_1, ..., \phi_p]'$—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 $\tilde{\sigma}^2=[(T\text{-}p\text{-}1)\hat{\sigma}^2_{OLS}]/\chi^2_{T-p-1}$, 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 $\tilde{\sigma}_j^2=[(T_j\text{-}p\text{-}1)\hat{\sigma}_j^2]/\chi^2_{T_j-p-1}$, 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 $\tau=t$ for one sided estimates, and $\tau=T$ 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 $\tau=t, T$.

Finally, for each t we take the mean across the 10,000 draws for both $\theta^i_{t|\tau}$ and $P^i_{t|\tau}$, $\tau=t, T$—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, $\gamma^k_{t|\tau} \equiv \mu^k_{t|\tau}/(1\text{-}\rho^k_{t|\tau})$. 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,[37] 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

[37] 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).

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 $k \to \infty$, converges to $V_\infty = 1+2\theta(1+\theta^2)^{-1}$, 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 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), Andrews (1993), and Bai (1997a): *exp*- and *sup*-Wald test statistics, and bootstrapped *p*-values

	exp-Wald	*sup*-Wald
U.S., output per hour:		
nonfarm business sector	1.09 (0.224)	5.28 (0.286)
business sector	2.51 (0.052)	9.16 (0.08)
nonfinancial corporations	4.44 (0.047)	14.74 (0.06)
manufacturing	2.96 (0.162)	10.86 (0.198)
manufacturing, durables	7.42 (0.055)	20.87 (0.063)
manufacturing, non durables	1.08 (0.384)	5.42 (0.375)
U.S., output per worker:		
nonfarm business sector	0.77 (0.370)	4.51 (0.404)
business sector	1.77 (0.118)	8.13 (0.130)
nonfinancial corporations	1.69 (0.275)	7.18 (0.292)
manufacturing	2.76 (0.249)	10.60 (0.246)
manufacturing, durables	3.76 (0.171)	13.10 (0.168)
manufacturing, non durables	0.44 (0.685)	2.53 (0.769)
U.K., real GDP:		
per hour	0.14 (0.923)	1.44 (0.917)
per worker	0.54 (0.708)	1.12 (0.889)
Eurozone, real GDP:		
per hour	9.41 (0.003)	25.72 (0.003)
per worker	16.91 (0)	42.51 (0.001)
Japan, real GDP:		
per hour	0.81 (0.348)	6.23 (0.247)
per worker	0.50 (0.598)	3.38 (0.579)
Australia:		
output per hour, market sector	0.35 (0.723)	2.11 (0.773)
output per hour, all industries	0.31 (0.54)	2.33 (0.554)
real GDP per worker	0.28 (0.852)	4.01 (0.588)

Table 2 Tests for multiple breaks in the mean based on Andrews and Ploberger (1994) and Bai (1997a): estimated break dates and 90% confidence intervals, and estimated mean productivity growth by sub-sample

Break dates and 90% confidence intervals	Sub-periods	Mean (standard error)
(a) Output per hour		
U.S., output per hour, business sector:		
1973:2 [1970:4; 1985:2]	1947:2-1973:1	3.27 (0.34)
1995:4 [1991:2; 2005:4]	1973:2-1995:3	1.44 (0.34)
	1995:4-2006:3	2.89 (0.25)
U.S., output per hour, nonfinancial corporations:		
1995:3 [1994:2; 2004:4]	1958:2-1995:2	1.89 (0.31)
	1995:3-2005:3	3.51 (0.28)
U.S., output per hour, manufacturing durables sector:		
1991:2 [1989:3; 1993:4]	1987:2-1991:1	1.41 (0.85)
	1991:2-2005:4	5.69 (0.46)
Eurozone, real GDP per hour:		
2001:1 [1973:1; 2003:1]	1980:2-2000:4	2.08 (0.18)
	2001:1-2005:4	0.61 (0.22)
(b) Output per worker		
Eurozone, real GDP per worker:		
1979:3 [1978:4; 1984:2]	1970:2-1979:2	3.33 (0.46)
	1979:3-2006:2	1.45 (0.17)

Table 3 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*
U.S., output per hour:		
nonfarm business sector	4.67 (0.401)	5.92 (0.326)
business sector	7.61 (0.089)	8.25 (0.084)
nonfinancial corporations	13.29 (0.091)	13.29 (0.128)
manufacturing	7.86 (0.309)	7.86 (0.395)
manufacturing, durables	17.81 (0.104)	17.81 (0.149)
manufacturing, non durables	3.09 (0.505)	3.09 (0.634)
U.S., output per worker:		
nonfarm business sector	4.97 (0.55)	5.74 (0.578)
business sector	8.19 (0.134)	8.19 (0.191)
nonfinancial corporations	10.34 (0.259)	11.59 (0.246)
manufacturing	12.33 (0.313)	12.33 (0.414)
manufacturing, durables	12.12 (0.377)	12.46 (0.457)
manufacturing, non durables	2.63 (0.806)	2.63 (0.885)
U.K., real GDP:		
per hour	2.35 (0.658)	3.17 (0.569)
per worker	3.68 (0.334)	3.89 (0.409)
Eurozone, real GDP:		
per hour	4.82 (0.083)	5.40 (0.08)
per worker	21.90 (0.007)	24.55 (0.004)
Japan, real GDP:		
per hour	3.84 (0.116)	3.84 (0.152)
per worker	6.98 (0.142)	7.82 (0.136)
Australia:		
output per hour, market sector	1.93 (0.884)	3.08 (0.747)
output per hour, all industries	2.47 (0.423)	3.95 (0.200)
real GDP per worker	4.31 (0.679)	4.83 (0.717)

Table 4 Tests for multiple breaks at unknown points in the sample in the mean based on Bai and Perron (1998): *sup*-F(ℓ+1$|\ell$) test statistics, and bootstrapped *p*-values				
	$F(2\|1)$	$F(3\|2)$	$F(4\|3)$	Estimated break dates and 90% confidence intervals
U.S., output per hour, business sector	8.20 (0.026)	1.88 (0.414)	0.30 (0.791)	1973:2 [1970:4; 1985:2] 1995:4 [1991:2; 2005:4]
Eurozone, real GDP:				
per hour	4.87 (0.043)	1.06 (0.495)	0.64 (0.467)	2001:1 [1973:1; 2003:1]
per worker	19.77 (0.002)	1.08 (0.84)	1.56 (0.362)	1977:1 [1976:2; 1982:3] 2000:3 [2000:2; 2002:1]

Table 5 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			
U.S., output per hour:	1947:2-1973:1	1973:2-1995:3	1995:4-2006:3
business sector	3.27 (0.34)	1.44 (0.34)	2.89 (0.25)
Eurozone, real GDP:			
per hour	1980:2-2000:4	2001:1-2005:4	
	2.09 (0.18)	0.64 (0.22)	
per worker	1970:2-1976:4	1977:1-2000:2	2000:3-2006:2
	3.64 (0.62)	1.82 (0.17)	0.55 (0.10)
Newey and West (1987) standard errors in parentheses.			

Table 6 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)	$\hat{\lambda}$
U.S., output per hour:				
nonfarm business sector	2.59 (0.105)	0.01897	12.88 (0.037)	0.02759
business sector	2.81 (0.073)	0.02069	11.49 (0.047)	0.02241
nonfinancial corporations	4.27 (0.021)	0.03276	13.96 (0.020)	0.03448
manufacturing sector	9.06 (0)	0.05172	24.85 (0)	0.05172
manufacturing sector, durables	9.99 (0)	0.05172	25.73 (0)	0.05172
manufacturing sector, non durables	3.66 (0.036)	0.05172	12.44 (0.033)	0.05172
U.S., output per worker:				
nonfarm business sector	1.35 (0.383)	0.00690	8.83 (0.159)	0.01724
business sector	2.51 (0.113)	0.01897	10.92 (0.062)	0.02241
nonfinancial corporations	2.96 (0.072)	0.02586	12.47 (0.034)	0.03103
manufacturing sector	4.51 (0.020)	0.05172	14.18 (0.024)	0.05172
manufacturing sector, durables	6.23 (0.006)	0.05172	17.55 (0.009)	0.05172
manufacturing sector, non durables	3.84 (0.032)	0.05172	12.21 (0.038)	0.05172
U.K., real GDP:				
per hour	0.63 (0.752)	0	3.60 (0.742)	0
per worker	0.65 (0.816)	0	4.71 (0.675)	0
Eurozone, real GDP:				
per hour	15.03 (0)	0.05172	36.99 (0)	0.05172
per worker	14.87 (0)	0.05172	38.12 (0)	0.05172
Japan, real GDP:				
per hour	1.117 (0.444)	0.01207	4.866 (0.490)	0.00345
per worker	1.190 (0.414)	0.01379	1.190 (0.414)	0.01379
Australia:				
output per hour, market sector	0.799 (0.621)	0	3.779 (0.678)	0
output per hour, all industries	2.817 (0.073)	0.04310	10.508 (0.067)	0.04655
real GDP per worker	0.725 (0.708)	0	0.725 (0.708)	0

Table 7 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, and 90% confidence interval
(a) Output per hour			
U.S., output per hour, nonfarm business sector:			
1983:3 [1973:2; 1993:4]	7.76 (0)	1947:2-1983:2	1.03 [8.9E-1; 1.21]
		1983:3-2006:3	0.39 [3.4E-1; 0.46]
United States, output per hour, business sector:			
1982:2 [1973:1; 1991:3]	9.53 (0)	1947:2-1982:1	1.11 [9.61E-1; 1.31]
		1982:2-2006:3	0.39 [3.4E-1; 0.46]
(b) Output per worker			
U.S., output per worker, nonfarm business sector:			
1983:4 [1976:2; 1991:2]	12.67 (0)	1947:2-1983:3	1.31 [1.13; 1.54]
		1983:4-2005:4	0.37 [3.2E-1; 0.44]
U.S., output per worker, business sector:			
1982:2 [1974:4; 1989:4]	13.54 (0)	1947:2-1982:1	1.36 [1.17; 1.60]
		1982:2-2005:4	0.38 [3.30; 0.45]
Eurozone, real GDP per worker:			
1992:3 [1986:3; 1998:3]	12.51 (0)	1970:2-1992:2	0.40 [3.41E-1; 0.46]
		1992:3-2006:2	0.08 [6.8E-2 ; 0.09]

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

Table 8 Power of the tests for breaks in the mean conditional on taking the estimated Stock-Watson TVP-MUB models as data generation processes

	Based on:	
	Andrews and Ploberger's (1994) *exp*-Wald test statistic[a]	Bai and Perron's (1998) *WDmax* test statistic
U.S., output per hour, nonfarm business sector	0.301	0.379
Australia, output per hour, all industries	0.374	0.434

[a] The test is for a single break at an unknown point in the sample.

Table 9 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	2.73	3.56	4.00	[2.20; 7.41]
business sector	1.84	2.70	3.19	[1.78; 6.34]
nonfinancial corporations	2.98	3.85	4.39	[2.47; 8.18]
		Output per worker:		
United States:				
nonfarm business sector	2.51	3.56	4.02	[2.20; 7.34]
business sector	2.00	3.07	3.50	[1.97; 6.53]
nonfinancial corporations	3.31	4.21	4.75	[2.70; 8.64]
United Kingdom, real GDP per worker	2.73	3.51	4.14	[2.13; 8.32]

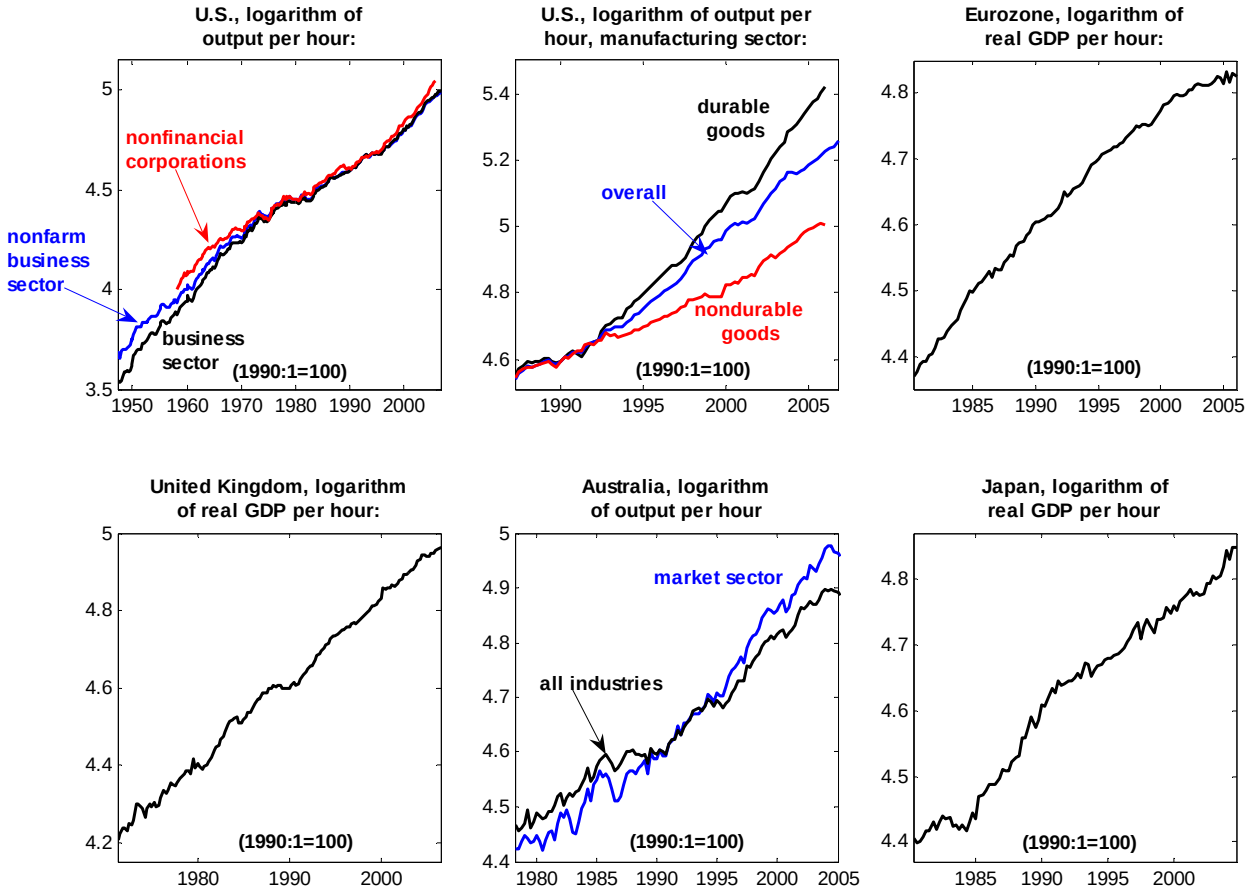

Figure 1 The raw data: logarithm of output per hour in the U.S., the U.K., the Eurozone, Australia, and Japan

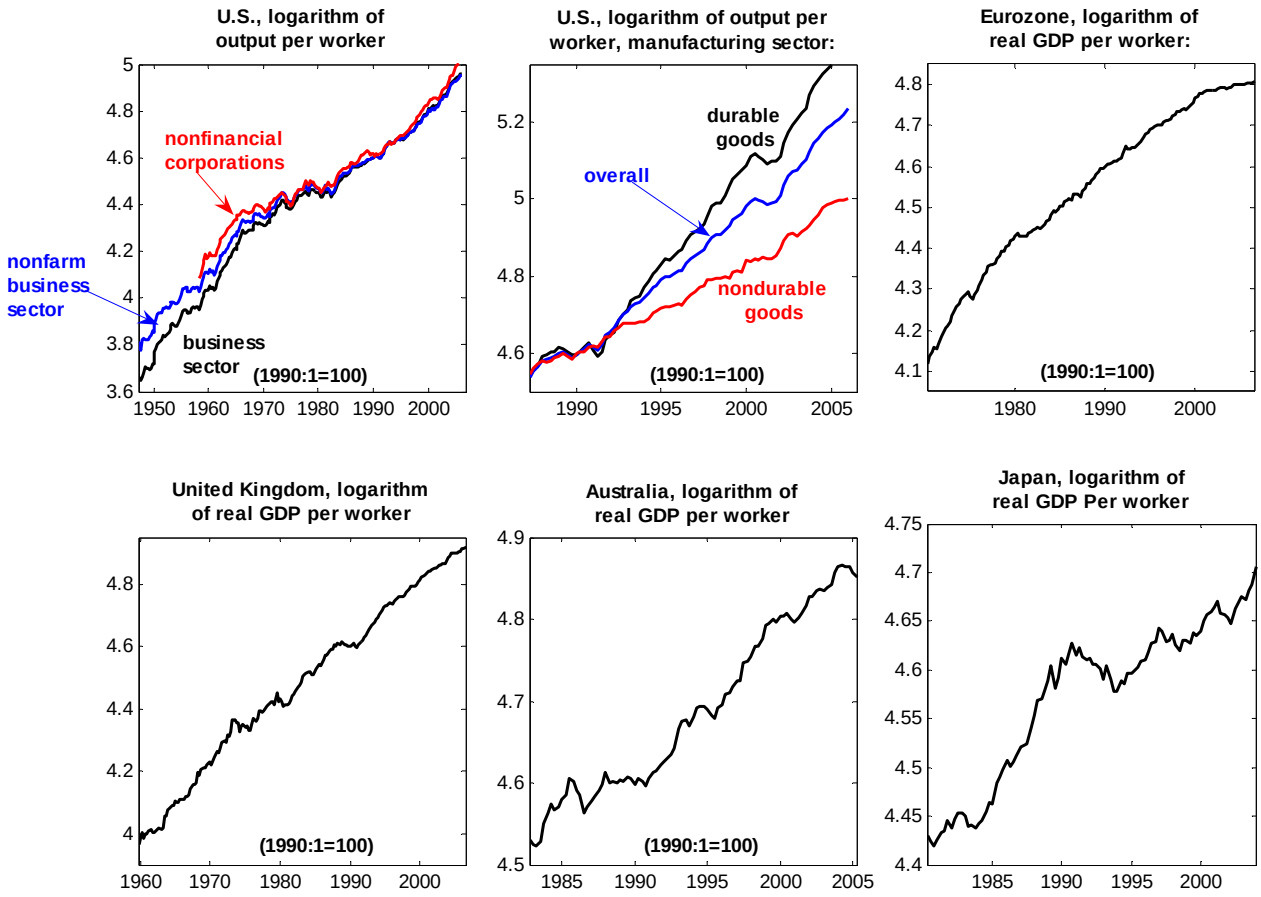

Figure 2 The raw data: logarithm of output per worker in the U.S., the U.K., the Eurozone, Australia, and Japan

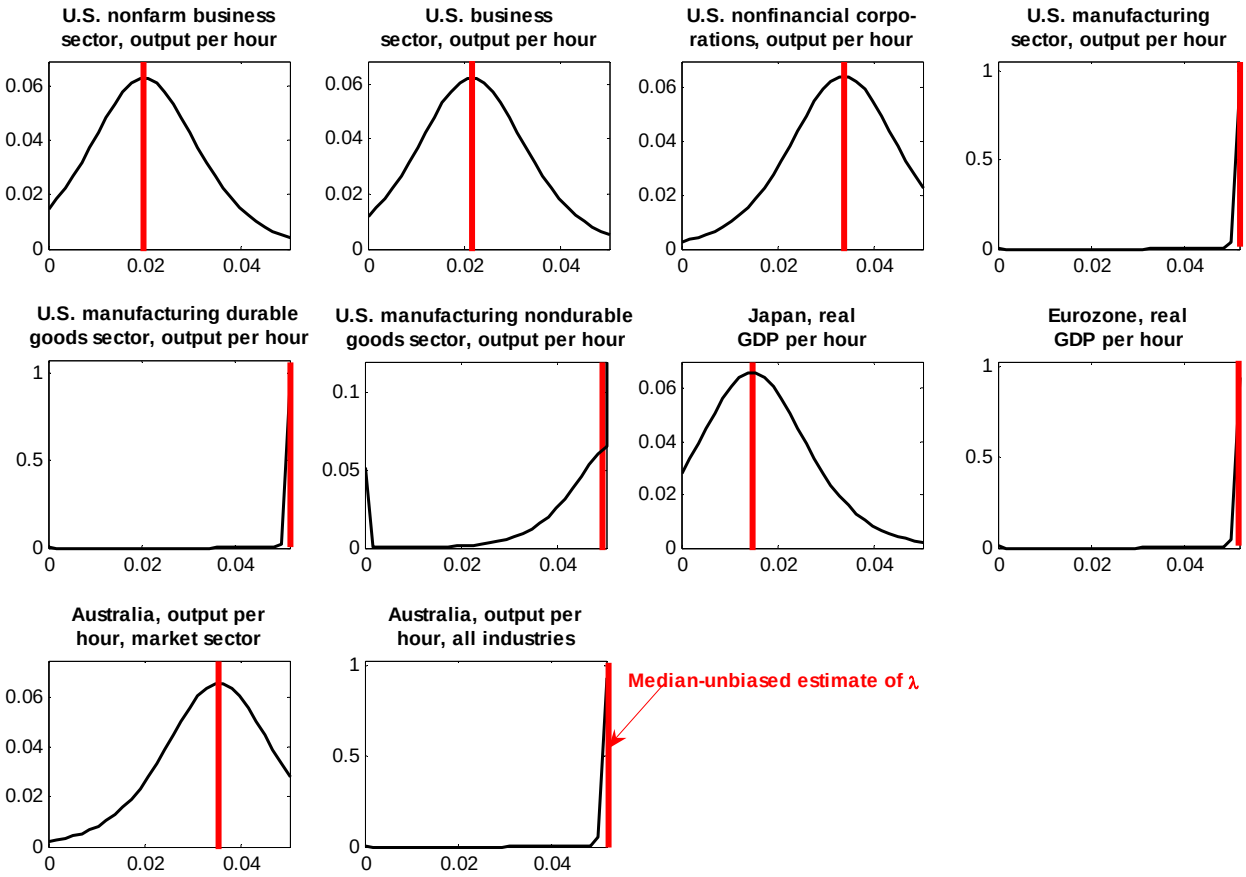

Figure 3 Median-unbiased estimates of λ, and deconvoluted PDFs of $\hat{\lambda}$ (output per hour)

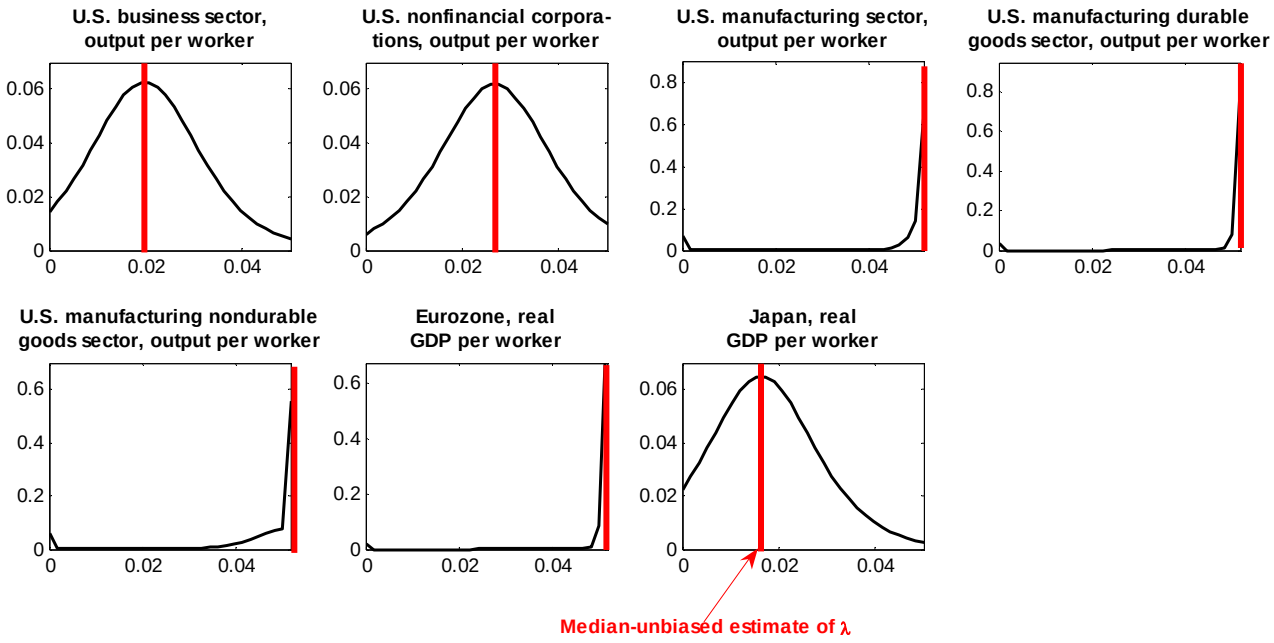

Figure 4 Median-unbiased estimates of λ, and deconvoluted PDFs of $\hat{\lambda}$ (output per worker)

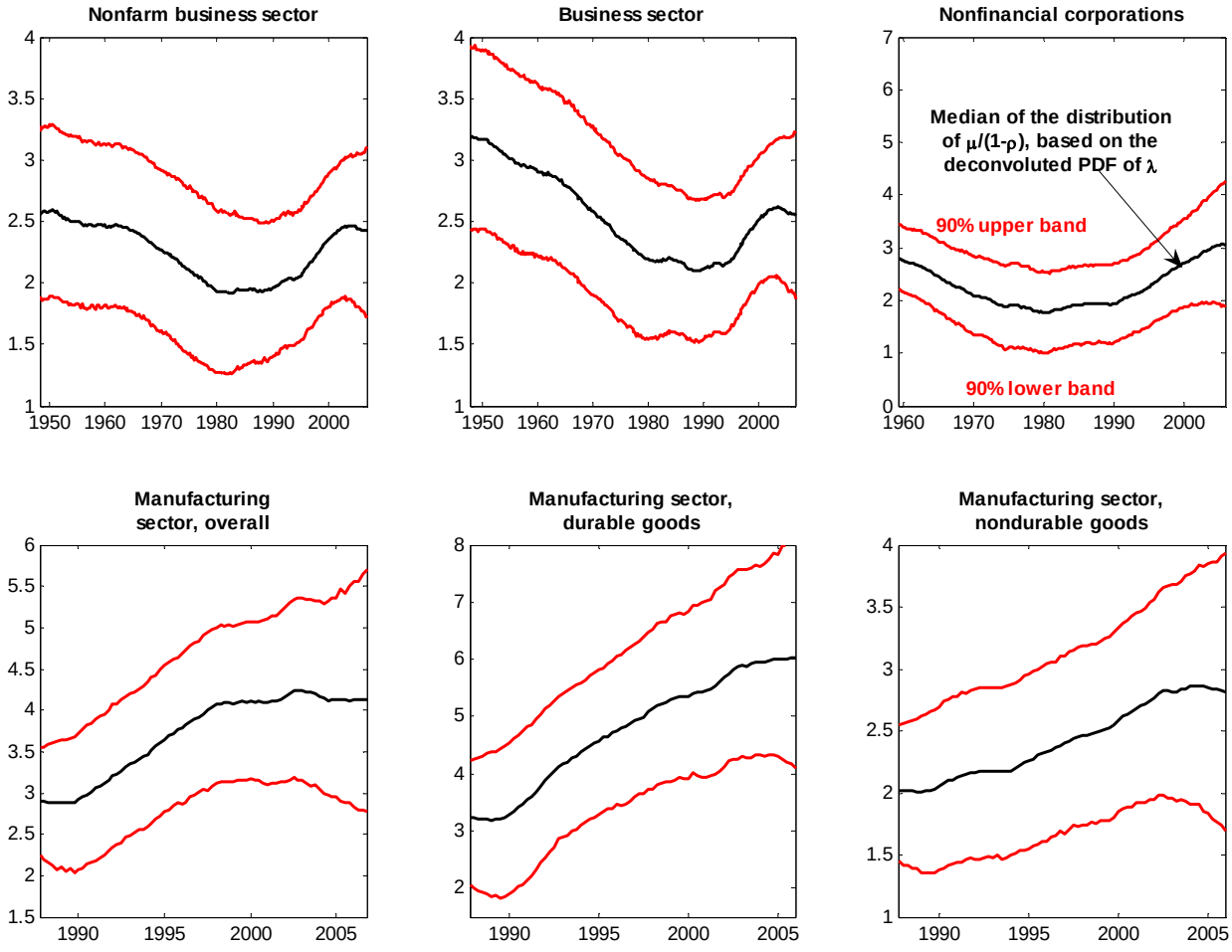

Figure 5 TVP-MUB two-sided estimates of equilibrium U.S. labor productivity growth, and 90% confidence bands (output per hour)

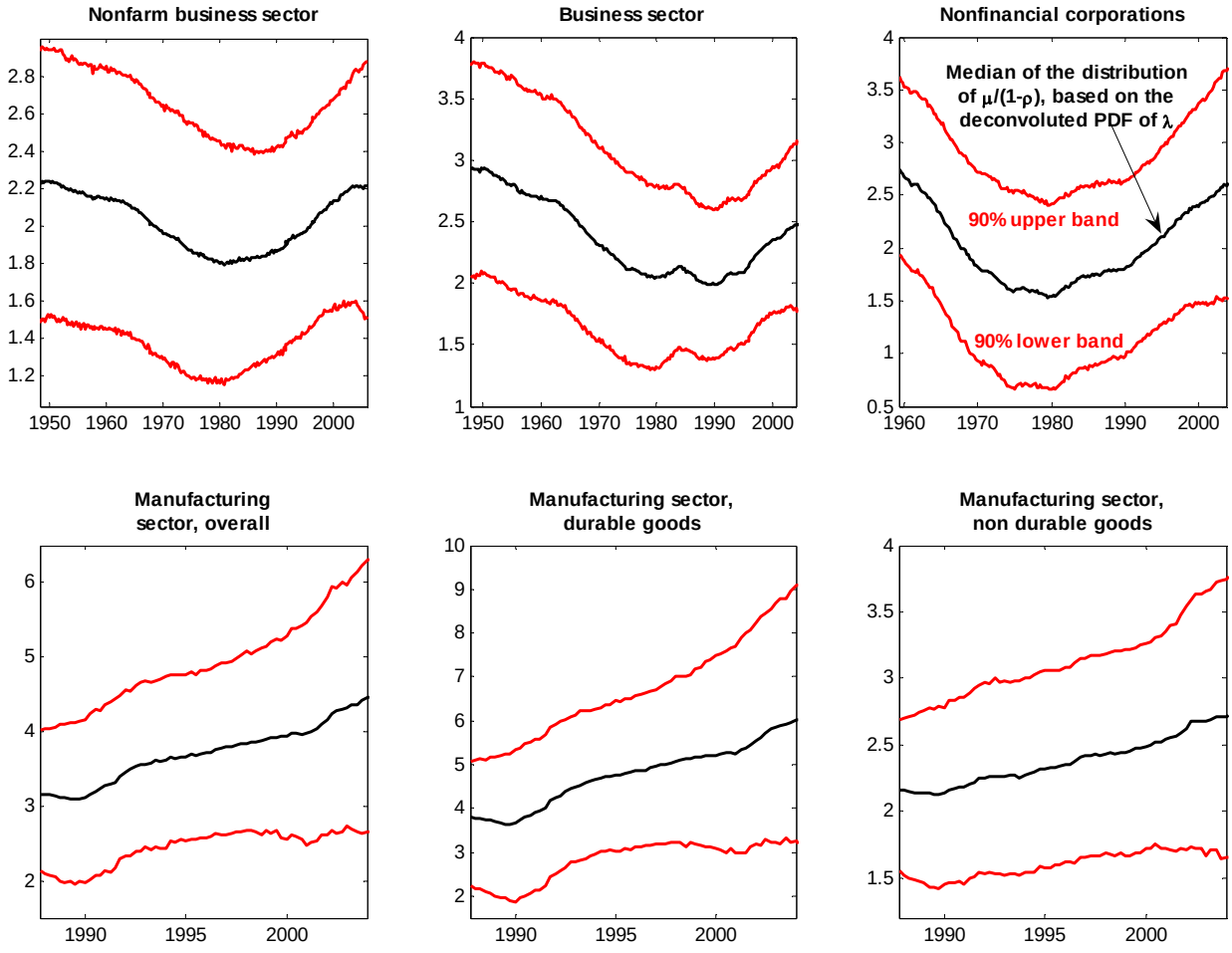

Figure 6 TVP-MUB two-sided estimates of equilibrium U.S. labor productivity growth, and 90% confidence bands (output per worker)

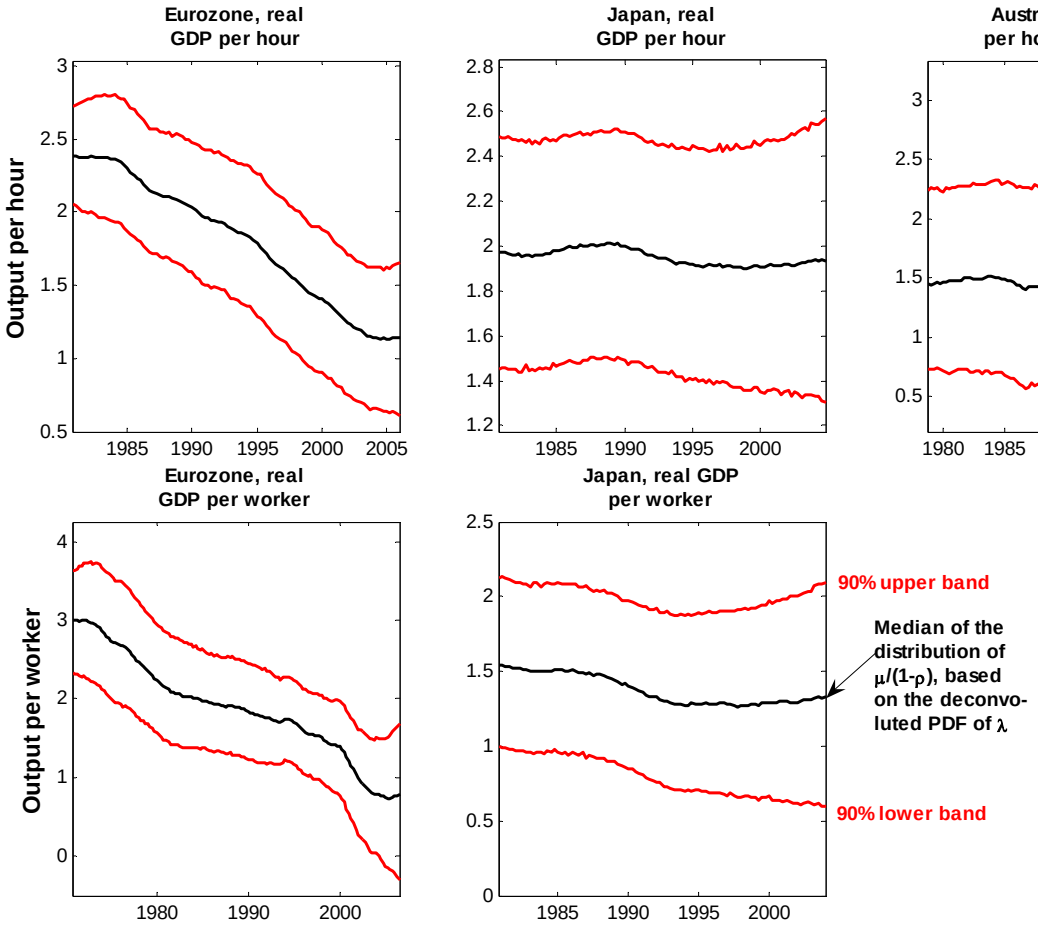

Figure 7 TVP-MUB two-sided estimates of equilibrium labor productivity growth in the Eurozone, Japan, and Australia, and 90% confidence bands (output per hour and output per worker)

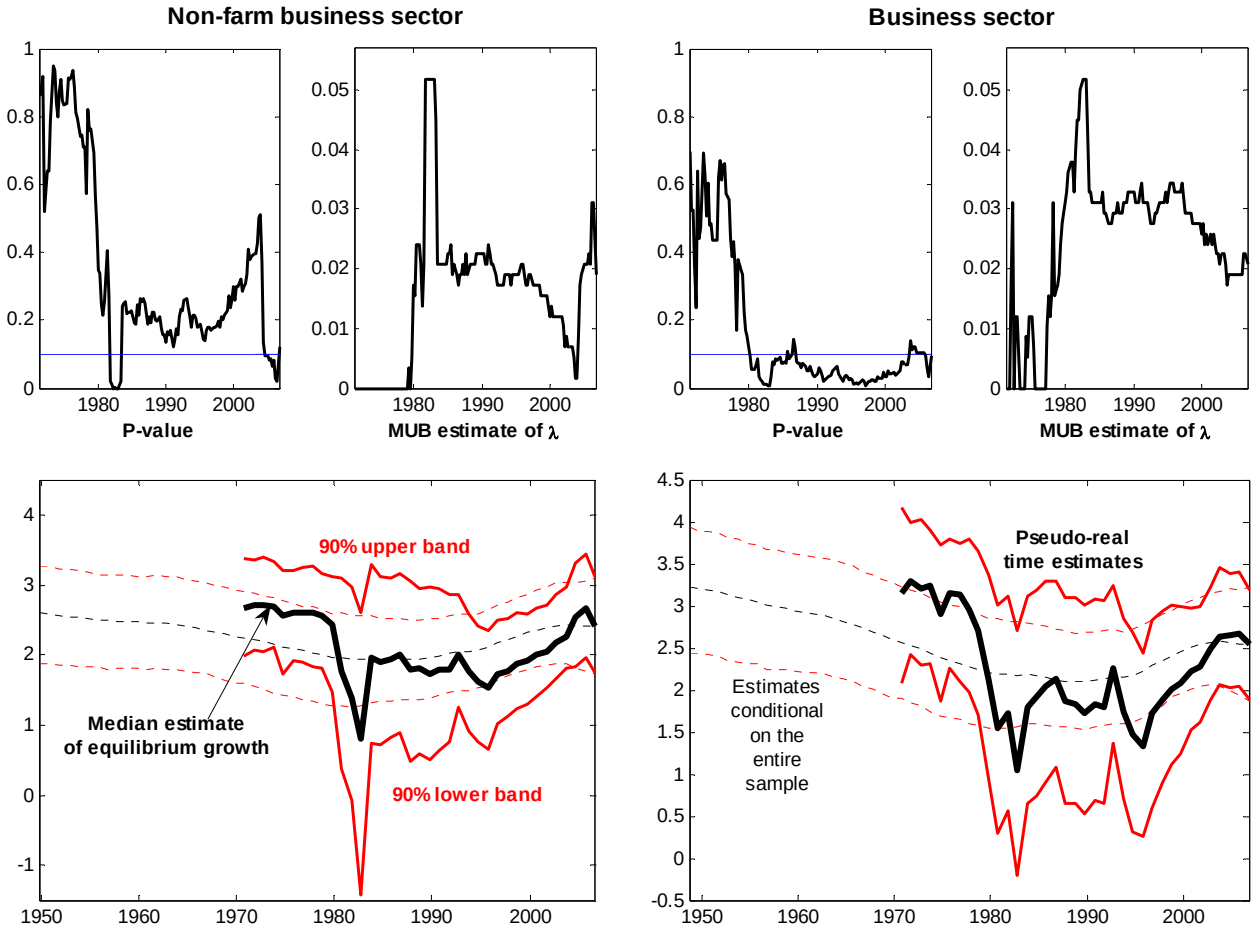

Figure 8 Pseudo-real time estimates of equilibrium productivity growth in the United States

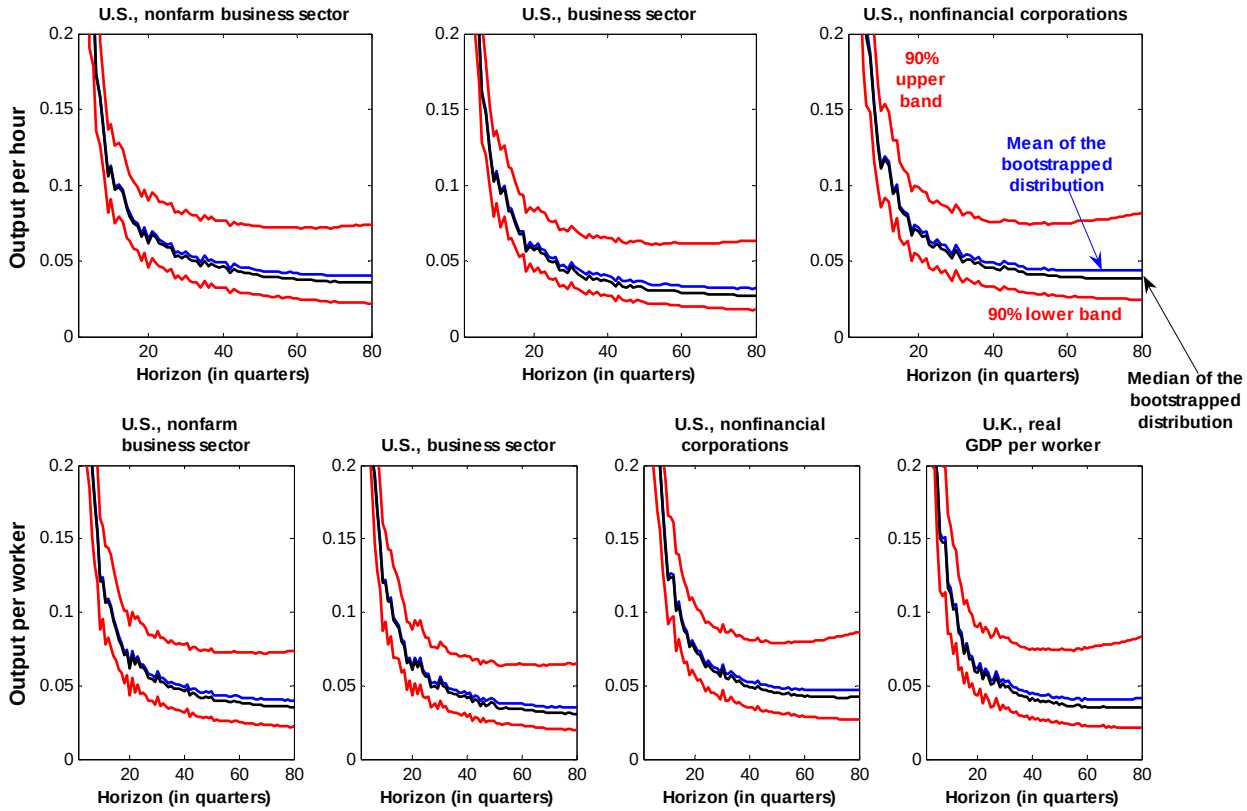

Figure 9 Estimates of the permanent component of the quarter-onquarter change in the rate of growth of labor productivity: results based on Cochrane s variance ratio estimator at the 20-year horizon

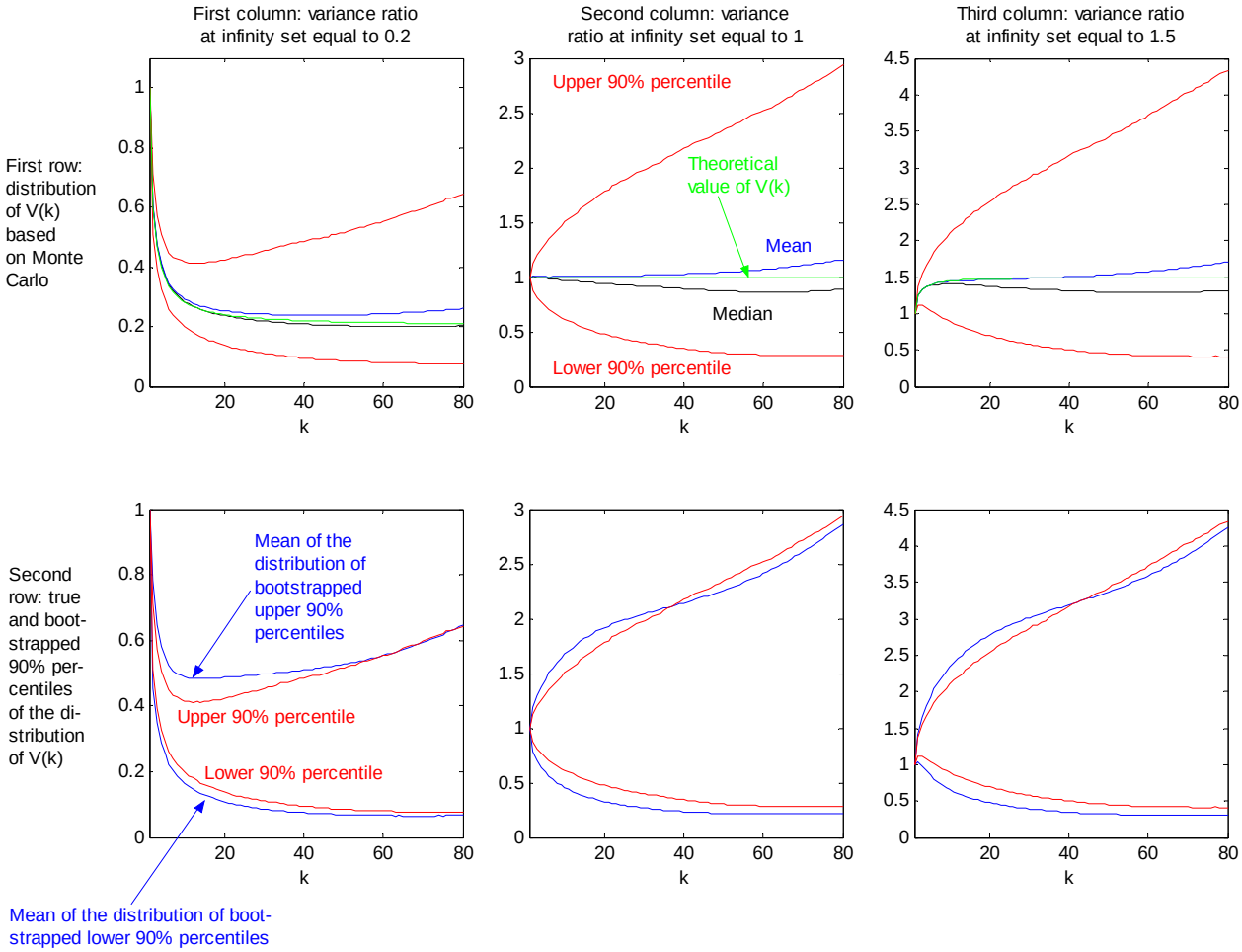

Figure 10 Monte Carlo evidence on the performance of Cochrane's (1988) variance ratio estimator, and on the accuracy of the confidence intervals based on spectral bootstrapping

European Central Bank Working Paper Series

For a complete list of Working Papers published by the ECB, please visit the ECB's website (http://www.ecb.int)

680 "Comparing alternative predictors based on large-panel factor models" by A. D'Agostino and D. Giannone, October 2006.

681 "Regional inflation dynamics within and across euro area countries and a comparison with the US" by G.W. Beck, K. Hubrich and M. Marcellino, October 2006.

682 "Is reversion to PPP in euro exchange rates non-linear?" by B. Schnatz, October 2006.

683 "Financial integration of new EU Member States" by L. Cappiello, B. Gérard, A. Kadareja and S. Manganelli, October 2006.

684 "Inflation dynamics and regime shifts" by J. Lendvai, October 2006.

685 "Home bias in global bond and equity markets: the role of real exchange rate volatility" by M. Fidora, M. Fratzscher and C. Thimann, October 2006

686 "Stale information, shocks and volatility" by R. Gropp and A. Kadareja, October 2006.

687 "Credit growth in Central and Eastern Europe: new (over)shooting stars?" by B. Égert, P. Backé and T. Zumer, October 2006.

688 "Determinants of workers' remittances: evidence from the European Neighbouring Region" by I. Schiopu and N. Siegfried, October 2006.

689 "The effect of financial development on the investment-cash flow relationship: cross-country evidence from Europe" by B. Becker and J. Sivadasan, October 2006.

690 "Optimal simple monetary policy rules and non-atomistic wage setters in a New-Keynesian framework" by S. Gnocchi, October 2006.

691 "The yield curve as a predictor and emerging economies" by A. Mehl, November 2006.

692 "Bayesian inference in cointegrated VAR models: with applications to the demand for euro area M3" by A. Warne, November 2006.

693 "Evaluating China's integration in world trade with a gravity model based benchmark" by M. Bussière and B. Schnatz, November 2006.

694 "Optimal currency shares in international reserves: the impact of the euro and the prospects for the dollar" by E. Papaioannou, R. Portes and G. Siourounis, November 2006.

695 "Geography or skills: What explains Fed watchers' forecast accuracy of US monetary policy?" by H. Berger, M. Ehrmann and M. Fratzscher, November 2006.

696 "What is global excess liquidity, and does it matter?" by R. Rüffer and L. Stracca, November 2006.

697 "How wages change: micro evidence from the International Wage Flexibility Project" by W.T. Dickens, L. Götte, E. L. Groshen, S. Holden, J. Messina, M. E. Schweitzer, J. Turunen, and M. E. Ward, November 2006.

698 "Optimal monetary policy rules with labor market frictions" by E. Faia, November 2006.

699 "The behaviour of producer prices: some evidence from the French PPI micro data" by E. Gautier, December 2006.

700 "Forecasting using a large number of predictors: Is Bayesian regression a valid alternative to principal components?" by C. De Mol, D. Giannone and L. Reichlin, December 2006.

701 "Is there a single frontier in a single European banking market?" by J.W. B. Bos and H. Schmiedel, December 2006.

702 "Comparing financial systems: a structural analysis" by S. Champonnois, December 2006.

703 "Comovements in volatility in the euro money market" by N. Cassola and C. Morana, December 2006.

704 "Are money and consumption additively separable in the euro area? A non-parametric approach" by B. E. Jones and L. Stracca, December 2006.

705 "What does a technology shock do? A VAR analysis with model-based sign restrictions" by L. Dedola and S. Neri, December 2006.

706 "What drives investors' behaviour in different FX market segments? A VAR-based return decomposition analysis" by O. Castrén, C. Osbat and M. Sydow, December 2006.

707 "Ramsey monetary policy with labour market frictions" by E. Faia, January 2007.

708 "Regional housing market spillovers in the US: lessons from regional divergences in a common monetary policy setting" by I. Vansteenkiste, January 2007.

709 "Quantifying and sustaining welfare gains from monetary commitment" by P. Levine, P. McAdam and J. Pearlman, January 2007.

710 "Pricing of settlement link services and mergers of central securities depositories" by J. Tapking, January 2007.

711 "What "hides" behind sovereign debt ratings?" by A. Afonso, P. Gomes and P. Rother, January 2007.

712 "Opening the black box: structural factor models with large cross-sections" by M. Forni, D. Giannone, M. Lippi and L. Reichlin, January 2007.

713 "Balance of payment crises in emerging markets: How early were the "early" warning signals?" by M. Bussière, January 2007.

714 "The dynamics of bank spreads and financial structure" by R. Gropp, C. Kok Sørensen and J.-D. Lichtenberger, January 2007.

715 "Emerging Asia's growth and integration: How autonomous are business cycles?" by R. Rüffer, M. Sánchez and J.-G. Shen, January 2007.

716 "Adjusting to the euro" by G. Fagan and V. Gaspar, January 2007.

717 "Discretion rather than rules? When is discretionary policy-making better than the timeless perspective?" by S. Sauer, January 2007.

718 "Drift and breaks in labor productivity" by L. Benati, January 2007.

ISSN 1561081-0