Jenaer Schriften zur Wirtschaftswissenschaft

On the Dynamics of the U.S. Manufacturing Productivity Distribution

Jens J. Krüger

5/2003

Arbeits- und Diskussionspapiere
der Wirtschaftswissenschaftlichen Fakultät
der Friedrich-Schiller-Universität Jena

ISSN 1611-1311

Herausgeber:

Wirtschaftswissenschaftliche Fakultät
Friedrich-Schiller-Universität Jena
Carl-Zeiß-Str. 3, 07743 Jena

www.wiwi.uni-jena.de

Schriftleitung:

Prof. Dr. Hans-Walter Lorenz
h.w.lorenz@wiwi.uni-jena.de

Prof. Dr. Armin Scholl
a.scholl@wiwi.uni-jena.de

On the Dynamics of the U.S. Manufacturing Productivity Distribution

by

Jens J. Krüger

Friedrich-Schiller-University Jena
February 2003

Department of Economics, Carl-Zeiß-Str. 3, D-07743 Jena, Germany
Tel.: +49 3641 943 203, Fax: +49 3641 943 202, E-Mail: jens.krueger@wiwi.uni-jena.de

Abstract:

In this paper a formal model of the productivity dynamics of manufacturing industries is developed with key features being the absence of optimal decisions and equilibrium coordination, heterogeneity of industries with respect to their innovative ability and cumulativeness of innovations together with the working of spillover effects. From that model the law of motion of the productivity distribution across the industries is derived and nonparametrically estimated using data for 140 three-digit U.S. manufacturing industries over the period 1958-96. The conclusion of a substantial role of persistence in the productivity development is further sharpened by the application of unit root and stationarity tests for panel data.

JEL classification: O30, L16, N10, C14, C23

Keywords: distribution dynamics, productivity, persistence, panel unit root test

1 Introduction

Most dynamic stochastic economic models focus on the law of motion that governs some kind of average development over time. Second moments like the variance are less frequently considered. Since we are hardly ever faced with representative entities when modelling the features of economic data, the dominant focus on average behavior is misplaced and more attention should be payed on the behavior of higher moments. At best, the whole distribution of the data and its changes over time should be the object of analysis.

In the present paper we are doing exactly this by looking at the dynamics of the distribution of total factor productivity in the industries of the U.S. manufacturing sector over the period 1958-96. We do this on the theoretical level by modelling distribution dynamics as a general Markov process which is later specialized to provide an illustration and to obtain more sharply testable hypotheses. This model is based on Loury's (1981) model of the dynamics of the earnings distribution over successive generations in which the individuals decide on the amount of training of their offspring that together with heterogeneous ability determines future earnings. In the empirical part nonparametric estimates of the productivity distribution and of the stochastic law of motion that controls the transitions to other regions of the distribution are presented. These are supplemented by an application of recently developed unit root and stationarity tests for panel data. All empirical findings lead to the conclusion that the productivity of industries is governed by a highly persistent stochastic process which implies that once gained productivity levels are unlikely to be quickly eroded but also that it is difficult to catch up from low productivity levels.

An alternative stochastic dynamic model in which changes of the distribution of economic activity are in the heart of the analysis is provided by Ericson and Pakes (1995). In this model heterogeneous firms with rational expectations decide about investment, entry and exit and the probability that a firm is in a certain (discrete) state of efficiency is governed by a Markov transition kernel. Firms solve dynamic programming problems over infinite time horizons that depend only on the current state and take the distributions of the outcomes of its competitors decisions as fixed. The decisions are coordinated by a Markov perfect Nash equilibrium in the sense of Maskin and Tirole (1988, 2001). Many features of this model like entry and exit are not relevant in the present context of differential productivity growth of industries. We also abandon equilibrium conditions and maximizing behavior of agents over long (in effect infinite) time horizons which are even less realistic on the industry level. In addition, the stochastic process of the distribution that results from our model is not ergodic in contrast to the model of Ericson and Pakes (1995). In this respect our model is related to that of Durlauf (1993), although the exact notion of nonergodicity is different there. The closest relation of our model exists to models of Gibrat-type growth processes which are surveyed by Sutton (1997) once the short-run dynamics have been eliminated from the analysis.

On the empirical side, there is a huge amount of evidence on the prevalence of differential growth of firms and industries assembled (see Caves (1998)). With respect to the growth of industries, Morrison and Siegel (1997) for example, obtain strong evidence of scale economies to human capital, research and development and information technology capital in two-digit U.S. manufacturing industries using a dynamic cost function approach. This gives support to endogenous growth models that often imply that per capita income differences between countries are persistent. The nonparametric empirical tools necessary to study distribution dynamics have been first applied by Quah (1996, 1997) in his research on the dynamics of the world income distribution, which has been found to become increasingly bimodal.

The paper proceeds as follows. In the following section 2 the modification of the Loury (1981) model is described under the heading general Markov model and the law of motion of the productivity density is derived. Section 3 restricts the general Markov model to the specification of the lognormal density for productivity and derives the dynamics of mean and variance which together determine the shape of the productivity density. After a brief description of the data and the nonparametric method for the calculation of total factor productivity in section 4 the results of the empirical analyses are reported in the following two sections. These empirical analyses consist of nonparametric estimates of the unconditional productivity densities and the transition density in section 5 which reveal information about the general Markov model. This is followed by a confirmatory analysis based on the joint application of unit root and stationarity tests for panel data in section 6 which is more closely related to the lognormal model. Finally, section 7 concludes with some interpretations and reservations.

2 General Markov Model

The general framework on which the model in the present paper is based is taken from Loury (1981). Loury models the intergenerational dynamics of the earnings distribution using a Markov process. In his model there is a continuum of utility maximizing individuals that constitute the mature generation and that produce an output depending on ability and the amount of training recieved. This output generates the income from which a part is spend for the training of the younger generation. The amount of training is an intergenerational money transfer which, together with the (fixed) ability distribution, translates the distribution of earnings from one generation to the next.

Here, the basic mechanism of the model is reinterpreted in a productivity context. It is assumed that there is a continuum of industries, each populated by a single firm which behaves myopically and does not exercise monopoly power. The productivity levels of the industries are distributed according to a certain distribution in each period *t*. Productivity allows (together with other input factors) to produce output that is sold on the different markets. Part of these sales is spend for R&D and related activities that, together with an index of ability, control the translation of the productivity distribution of period *t* into the productivity distribution of period *t*+1.

Formally, the law of motion for the productivity level of each single industry can be stated as

$$a_{t+1} = h(\varphi, e(y(a_t))),\ \ t = 1, 2, ...,$$

where $a_t \in \mathbb{R}_+$ denotes the productivity level that generates sales y according to $y(a_t) \in \mathbb{R}_+,\ y'(a_t) > 0$. The input into the innovation process (e.g. R&D effort) depending on y is denoted by $e(y) \in \mathbb{R}_+, e'(y) > 0$ and $\varphi \in \mathbb{R}_+$ is an index of innovative (and entrepreneurial) ability with density $f(\varphi)I_{[0,\infty)}(\varphi)$, where I represents the usual indicator function. The function $h(\cdot,\cdot)$ is increasing in both arguments and specifies how the research input and the ability index together affect next periods productivity level a_{t+1}.

The probability space of the productivity level of an industry in period t, a_t, is defined by the triple $\{\Omega, \Re, v_t\}$, where the sample space is $\Omega = \mathbb{R}_+$ (the set of non-negative real numbers), $\Re$

is the σ-field generated by the one-dimensional Borel sets and v_t is a probability measure defined on $\Re$. Thus, the productivity distribution of the whole manufacturing sector in period t is represented by the probability measure v_t. This probability measure is transformed into the next periods probability measure v_{t+1} by a Markov process that generates a sequence of probability measures $\{v_t\}$ according to the law of motion

$$v_{t+1}(A) = \int_{\Omega} P(e(y(x)), A) v_t(dx)\ ,\ t = 1, 2, ...,$$

where the transition function (or transition kernel) $P(e(y(x)), A) = \Pr(a_{t+1} \in A | a_t = x)$ gives the probability that next periods productivity level is in the set A, given a current period productivity level of x which leads to sales y that enable research expenditures e. $P : \Omega \times \Re \rightarrow [0, 1]$ is a nonnegative function that satisfies the requirements of being a $\Re$-measurable function in the first argument and a probability measure in the second argument (see Feller (1971, p. 205), Stokey/Lucas (1989, ch. 8) or Durlauf/Quah (1998, p. 300)). The integration amounts to a weighted summation of the transition probabilities over all possible current productivity levels. Thus, the transition kernel is a complete probabilistic description of the transitions from a state x to any other portion of the state space Ω.

The connection of the law of motion of $\{a_t\}$ to the law of motion of $\{v_t\}$ is established through the derivation of the transition kernel P by

$$P(e(y(x)), A) = \int_{h^{-1}(A, e(y(x)))} f(\varphi) \mu(d\varphi),$$

where integration is with respect to Lebesgue measure μ over the set of abilities that lead to a next periods productivity level in the set A, given that the current period R&D effort is e, that is the set $h^{-1}(A, e) = \{\varphi \in [0, \infty) : h(\varphi, e) \in A\}$. $P(\cdot, A)$ is measurable if the functions $y(\cdot)$ and $e(\cdot)$ are both continuous.

Since it is much more convenient to work with density functions than with measures on abstract probability spaces the next step is to derive the law of motion of the density functions from the law of motion of $\{v_t\}$. First, let $g_t(\cdot)$ be the Radom-Nikodym derivative of v_t with respect to Lebesgue measure (see Billingsley (1995, sect. 32)) which exists and is unique if v_t is absolutely continuous with respect to Lebesgue measure μ (denoted $v_t \ll \mu$). It follows that the measure v_t has a density function $g_t(\cdot)$ with respect to Lebesgue measure, that is $v_t(X) = \int_X g_t(x)\mu(dx) = \int_X g_t(x)dx$ for $X \in \Re$ and we can state $v_t(dx) = dv_t(x) = g_t(x)dx$. Second, $[0, a] \in \Re$ for $a \in$ ‘$_+$ fixed implies that $v_{t+1}([0, a]) = \int_\Omega P(e(y(x)), [0, a])v_t(dx)$. Since $\Omega =$ ‘$_+$, $v_{t+1}([0, a])$ is simply a distribution function on ‘$_+$, denoted by $G_{t+1}(a)$. It follows that the Markov process can now be expressed as $G_{t+1}(a) = \int_\Omega P(e(y(x)), [0, a])g_t(x)dx$. To reach a formulation that is a mapping from a density function to a density function we differentiate $G_{t+1}(a)$ with respect to a and obtain

$$g_{t+1}(a) = \frac{dG_{t+1}(a)}{da} = \int_\Omega \frac{\partial P(e(y(x)),[0,a])}{\partial a} g_t(x)dx = \int_\Omega p(e(y(x)), a)g_t(x)dx,$$

where sufficient regularity to permit the exchange of integration and differentiation is supposed. The function $p(e(y(x)), a) = \partial P(e(y(x)), [0, a])/\partial a$ is called a transition density (or stochastic density kernel; see Feller (1971, p. 205)). In sum, the law of motion of the density is

$$g_{t+1}(a_{t+1}) = \int_\Omega p(e(y(a_t)), a_{t+1})g_t(a_t)da_t.$$

In the next section more structure is put on this general framework model. It is important to note that only very few functional forms for the ability distribution $f(\cdot)$ and the law of motion of the productivity level $h(\cdot, \cdot)$ exist from which the squence of density functions $\{g_t\}$ can be derived analytically. Since these functional forms are quite restrictive, future research will focus on numerical simulations of the model using less restrictive specifications than the one presented below.

3 Lognormal Specification

To illustrate the general model outlined above we specify the real sales of each industry through a Cobb-Douglas functional form that assumes that real sales depend on the industry's own technology level a_t and the mean technology level $\bar{a}_t = \mathrm{E}(a_t)$ in the manufacturing sector

$$y(a_t) = \gamma_1 a_t^{\gamma_2} \bar{a}_t^{\gamma_3}, \gamma_1, \gamma_2 > 0.$$

The parameter γ_3 captures the balance between spillover-effects that influence the sales of the industry positively and competitive pressure that exerts an opposite effect on the sales. If $\gamma_3 > 0$ then the spillover-effects outweigh the effects from competitive pressure, otherwise we should expect $\gamma_3 < 0$. Suppose that each industry is represented by a single firm that acts boundedly rational in that it invests a constant share $\psi \in [0, 1]$ of sales in innovative activities, that is $e(y(a_t)) = \psi \cdot y(a_t)$. Plugging in the sales function gives $e(y(a_t)) = \psi\gamma_1 a_t^{\gamma_2} \bar{a}_t^{\gamma_3}$. It is assumed that the ability index φ and the amount of investment in innovative activities interact multiplicatively $h(\varphi, e_t) = \varphi \cdot e(y(a_t))$ to determine the productivity level of the next period, so that the law of motion is

$$a_{t+1} = h(\varphi, e(y(a_t))) = \varphi\psi\gamma_1 a_t^{\gamma_2} \bar{a}_t^{\gamma_3}.$$

This specification comprises mechanisms of endogenous growth models (e.g. Romer (1986) or Jones (1995)) dependending on the parameter values and also contains a time-to-build aspect in that current investment affects only next periods productivity level. Since ψ and φ are interacted multiplicatively, differences in ability can alternatively be interpreted as differences in the propensity to invest in innovative activities.

The model departs from the usual neoclassical modelling in several respects. First, the different industries are heterogeneous with respect to their innovative ability which influences their differential innovative success. Second, each industry invests a constant share of sales in innovative activities and does not follow an optimal investment path that schedules investment into the remote future. Third, the productivity development of each industry is an open ended process that evolves without being constrained by any forces that drive the process towards an equilibrium growth path. This occurs despite the fact that the mean of this process shows very regular behavior, as will be derived below.

The basic idea of the parametric specification of the distributional side of the model is to use the property of the bivariate normal distribution that all marginal and conditional distributions are also of the normal family. But since normal distributions have support on the whole ‘ and not just ‘$_+$, we have to express the relevant variables in logs. This implies for the general Markov model of the preceding section a new probability space with $\Omega =$ ‘ and the corresponding σ-field is generated by the Borel sets on which the sequence of probability measures $\{\nu_t\}$ is defined. Despite this change all measure theoretic considerations above that lead to the derivation of the transition density $p(\cdot,\cdot)$ go through on this new probability space word for word.

Denoting the logs of the functions $y(\cdot)$, $e(\cdot)$ and $h(\cdot,\cdot)$ by $\tilde{y}(\cdot)$, $\tilde{e}(\cdot)$ and $\tilde{h}(\cdot,\cdot)$, the set of values for $\ln\varphi$ that lead to a next periods log productivity level in the set $A = (-\infty, \ln a_{t+1}]$ given the log amount of investment $\tilde{e}(\tilde{y}(\ln a_t))$ is $\tilde{h}^{-1}(A, \tilde{e}(\tilde{y}(\ln a_t))) = \{\ln\varphi \in ‘ : \tilde{h}(\ln\varphi, \tilde{e}(\tilde{y}(\ln a_t))) \in A\}$. Using the continuity of $\ln a_{t+1} = \tilde{h}(\ln\varphi, \tilde{e}(\tilde{y}(\ln a_t))) = \ln\varphi + \ln(\psi\gamma_1) + \gamma_2 \ln a_t + \gamma_3 \ln \bar{a}_t$ it can be stated that

$$\ln\varphi \in \tilde{h}^{-1}(A, \tilde{e}(\tilde{y}(\ln a_t))) = (-\infty, \ln a_{t+1} - \ln(\psi\gamma_1) - \gamma_2 \ln a_t - \gamma_3 \ln \bar{a}_t].$$

The essential distributional assumption is now that $\ln\varphi \sim N(0, \sigma_\varphi^2)$ from which the functional form of the transition kernel follows immediately as

$$\begin{aligned}\tilde{P}(\tilde{e}(\tilde{y}(\ln a_t)), (-\infty, \ln a_{t+1}]) &= \int_{h^{-1}((-\infty,\ln a_{t+1}],\tilde{e}(\tilde{y}(\ln a_t)))} \phi(\sigma_\varphi^{-1}\ln\varphi)\, d\ln\varphi \\ &= \int_{-\infty}^{\ln a_{t+1}-\ln(\psi\gamma_1)-\gamma_2 \ln a_t-\gamma_3 \ln\bar{a}_t} \phi(\sigma_\varphi^{-1}\ln\varphi)\, d\ln\varphi \\ &= \Phi(\sigma_\varphi^{-1}(\ln a_{t+1} - \ln(\psi\gamma_1) - \gamma_2 \ln a_t - \gamma_3 \ln\bar{a}_t)),\end{aligned}$$

where $\Phi(\cdot)$ is the cumulative distribution function and $\phi(\cdot)$ is the density function of the standard normal distribution. This result fulfills the requirements of the definition of a transition function in that it is a distribution function (and therefore a probability measure) in $\ln a_{t+1}$ and a measurable function with respect to the Borel sets in $\ln a_t$ as a consequence of continuity. The corresponding density kernel is obtained through differentiation as

$$\begin{aligned}\tilde{p}(\tilde{e}(\tilde{y}(\ln a_t)), \ln a_{t+1}) &= \frac{\partial\Phi(\sigma_\varphi^{-1}(\ln a_{t+1} - \ln(\psi\gamma_1) - \gamma_2 \ln a_t - \gamma_3 \ln\bar{a}_t))}{\partial \ln a_{t+1}} \\ &= \phi(\sigma_\varphi^{-1}(\ln a_{t+1} - \ln(\psi\gamma_1) - \gamma_2 \ln a_t - \gamma_3 \ln\bar{a}_t)) \cdot \sigma_\varphi^{-1} \\ &= \frac{1}{\sqrt{2\pi\sigma_\varphi^2}} \exp(-\tfrac{1}{2\sigma_\varphi^2}(\ln a_{t+1} - \ln(\psi\gamma_1) - \gamma_2 \ln a_t - \gamma_3 \ln\bar{a}_t)^2).\end{aligned}$$

Since the normal distribution pertains to an exponential family, the exchange of integration and differentiation is permitted by theorem 5.8 of Lehmann and Casella (1998).

To complete the specification of the model we need an initial condition for the density of $\ln a_t$. Assuming that $\ln a_t \sim N(\mu_t, \sigma_t^2)$ implies that a_t is lognormally distributed with a mean of $\mathrm{E}(a_t) = e^{\mu_t + \sigma_t^2/2}$ from which $\ln \bar{a}_t$ can be expressed as $\mu_t + \sigma_t^2/2$. Using the just derived stochastic density kernel which represents the condiditional distribution function of $\ln a_{t+1}$ given $\ln a_t$

$$\ln a_{t+1} | \ln a_t \sim N(\ln(\psi\gamma_1) + (\gamma_2 + \gamma_3)\mu_t + \gamma_3\sigma_t^2/2 + \gamma_2(\ln a_t - \mu_t), \sigma_\varphi^2),$$

where the mean is stated in a form that facilitates the statement of the density of $\ln a_{t+1}$ (which is also normal) from the standard properties of the bivariate normal distribution. The expressions for the mean and variance of the conditional normal distribution (see Gallant (1997, p. 115)) can be solved for the mean and variance of the distribution of $\ln a_{t+1}$ which are given by

$$\mu_{t+1} = \ln(\psi\gamma_1) + (\gamma_2 + \gamma_3)\mu_t + \gamma_3\sigma_t^2/2 \quad \text{and} \quad \sigma_{t+1}^2 = \sigma_\varphi^2 + \gamma_2^2\sigma_t^2.$$

Thus, mean and variance of the log manufacturing productivity distribution evolve as linear processes with drift. The higher the dependence of the log productivity level of one period to another is on average the faster increases the mean log productivity level and its dispersion. In addition, the larger is the effect of spillovers the faster is the growth of the mean, depending on both mean and variance of the distribution of the previous period log productivity. Thus the stochastic process in this model never settles to a ergodic limiting distribution. This has to be distinguished from other types of nonergodicity such as that featured by Durlauf (1993) which is related to a class of conditional probability measures that is consistent with a multiplicity of invariant probability measures. In Arthur's (1989) analysis historical small events may have a decisive effect on the long-run outcome of the process of technology adoption. This leads to a path-dependent process which is also inconsistent with the notion of ergodicity.

These moments allow us to compute the distribution of log productivity of period $t + 1$ as $\ln a_{t+1} \sim N(\mu_{t+1}, \sigma_{t+1}^2)$ with density function $g_{t+1}(\ln a_{t+1})$ from the knowledge of previous period moments and model parameters. It follows for the distribution of productivity that $a_{t+1} \sim LN(\mu_{t+1}, \sigma_{t+1}^2)$ where LN denotes the lognormal distribution which has the density function

$$g_{t+1}(a_{t+1}) = \frac{1}{a_{t+1}\sqrt{2\pi\sigma_{t+1}^2}} \exp(-\tfrac{1}{2\sigma_{t+1}^2}(\ln a_{t+1} - \mu_{t+1})^2).$$

The moments of this distribution are related to the normal distribution through the parameters μ_{t+1} and σ_{t+1}^2. Johnson and Kotz (1970, pp. 115f.) assemble the formulas for the expectation $\mathrm{E}(a_{t+1}) = e^{\mu_{t+1}+\sigma_{t+1}^2/2}$, the variance $\mathrm{Var}(a_{t+1}) = e^{2\mu_{t+1}+\sigma_{t+1}^2}(e^{\sigma_{t+1}^2} - 1)$, the skewness coefficient $\mathrm{skewcoeff}(a_{t+1}) = \zeta_{t+1}^3 + 3\zeta_{t+1}$ (where $\zeta_{t+1} = \sqrt{e^{\sigma_{t+1}^2} - 1}$), and the kurtosis coefficient $\mathrm{kurtcoeff}(a_{t+1}) = \zeta_{t+1}^8 + 6\zeta_{t+1}^6 + 15\zeta_{t+1}^4 + 16\zeta_{t+1}^2$ of a lognormally distributed random variable.

In the following sections the results of a number of empirical analyses are reported. These comprise direct nonparametric kernel estimates of the sequence of density functions $\{g_t(a_t)\}$ and the transition density function $p(a_t, a_{t+1})$ using data for 140 three-digit SIC U.S. manufacturing industries. Since productivity is always positive it is preferable to estimate the density functions $\{\tilde{g}_t(\ln a_t)\}$ and the transition density function $\tilde{p}(\ln a_t, \ln a_{t+1})$ for log productivity in order to avoid boundary bias problems in the case of the kernel estimates. Since productivity and log productivity are related by a monotonically increasing function the interpretations on the log scale are directly transferable to the original scale of productivity. The subsequent application of panel unit root tests to these data sharpens some of the conclusions that we have obtained from the nonparametric estimates and allows for inference regarding the magnitude of the parameter γ_2 which determines the amount of persistence of the productivity differences between the manufacturing industries. Before we turn to the results we briefly discuss the data and the method that is used to calculate total factor productivity.

4 Data and Productivity Scores

The data used to calculate the productivity scores are from the NBER-CES manufacturing industry database which is described by Bartelsman and Gray (1996). This unique database provides consistent yearly time series over the period 1958-96 for quantity and price data of 459 manufacturing industries on the four-digit level. The nonparametric productivity measurement is performed with the following specification of the output variable and the six labor, capital and material/energy input variables (the abbreviations in square brackets refer to those defined in the data appendix of Bartelsman and Gray (1996)):

- output:
  real value of shipments [VSHIP/PISHIP]

- labor inputs:
  number of non-production workers [EMP–PRODE]
  production worker hours [PRODH]

- capital inputs:
  real equipment capital stock [EQUIP]
  real structures capital stock [PLANT]

- material and energy:
  real cost of materials other than electricity and fuels [(MATCOST–ENERGY)/PIMAT]
  real expenditures on fuels and electricity [ENERGY/PIEN]

The variables are aggregated from the four-digit to the three-digit level, which results in 140 three-digit industries.[1] Using these data the productivity scores are calculated by the Andersen-Petersen variant of data envelopment analysis (see Andersen and Petersen (1993)). Data envelopment analysis is a nonparametric method to calculate an index of total factor productivity by the distance of the input-output combinations of the industries from a data-determined piece-wise linear frontier production function. This method requires only quantity data of any desired scaling and is thoroughly surveyed in Charnes et al. (1994). The Andersen-Petersen model calculates productivity by calculating an index that indicates on which level the output of an industry has to be increased in order to reach a facet of the (piece-wise linear) frontier production function that is determined by the observations of the other $N-1$ industries, excluding the one for which efficiency is evaluated. Formally, the score for total factor productivity of industry i in year t, a_{it}, is the inverse of the quantity ϕ_{it} that is the solution of the following linear programming problem

$$\max\left\{\phi_{it} : \phi_{it} y_{it} \leq \sum_{h \in \{1,\dots,N\}\backslash i} \lambda_h y_{ht} \; ; \sum_{h \in \{1,\dots,N\}\backslash i} \lambda_h \boldsymbol{x}_{ht} \leq \boldsymbol{x}_{it} \; ; \boldsymbol{\lambda}_{-i} \geq 0\right\}$$

where y_{it} is the output variable specified above, $\boldsymbol{x}_{it}$ is a 6×1-vector that comprises the observations for the six input variables of industry i in year t and $\boldsymbol{\lambda}_{-i}$ denotes the $(N–1)$-vector of the λ's omitting the i-th element.[2]

[1] SIC 3292 (asbestos) has been removed from the data base because of missing data after 1993.
[2] This procedure is completely deterministic. There exists an alternative econometric approach to the estimation of fontier functions that promises to be able to divide measurement error from the productivity measure (see

The calculations of the Andersen-Petersen model for each year separately results in a balanced panel of total factor productivity scores a_{it} $(i = 1, ..., N,\ t = 1, ..., T)$ that variate between zero and infinity for $N = 140$ three-digit industries over the $T = 39$ years covering the period 1958-96. The majority of the observations that do not determine the frontier function get assigned a productivity score smaller than unity. Those observations that determine the frontier function get assigned a productivity score larger than unity. To avoid boundary problems in the case of the nonparametric kernel density estimates all subsequent empirical analyses are based on log productivity scores $\ln a_{it}$ which are allowed to variate on the whole ‘ .

5 Nonparametric Estimation

Turning to the empirical results we first discuss the nonparametric estimates of the sequence of density functions $\{\tilde{g}_t(\ln a_t)\}$ before we turn to the transition density function $\tilde{p}(\ln a_t, \ln a_{t+1})$ of the 140 three-digit U.S. manufacturing industries. The latter is much more important since it provides a multitude of information concerning the productivity dynamics, much more than density functions as unconditional objects are able to do.

The density function $\tilde{g}_t(\ln a_t)$ for each period are nonparametrically estimated by the univariate kernel density estimator for the density at point x (see Wand and Jones (1995, ch. 2))

$$\hat{g}_t(x) = \frac{1}{Nh_t} \sum_{i=1}^{N} K\left(\frac{x - \ln a_{it}}{h_t}\right),$$

where the standard normal density is used as kernel function $K(\cdot)$ and the bandwidth h_t is chosen by the Sheather-Jones 2nd generation bandwidth estimator (Sheather and Jones (1991)) for each period separately. This bandwidth estimator has proved to be the prefered method for one-dimensional kernel density estimation in the comparison of Jones et al. (1996).

Figure 1 shows the kernel density estimates of the manufacturing productivity distribution for the years 1958, 1969, 1985 and 1996 which are chosen with regard to Jorgenson (1990) and provide an approximately equidistant subdivision of the sample period.

e.g. Greene 1993). However, the Monte Carlo studies of Banker et al. (1993) and Ruggiero (1999) show that this advantage of the econometric appoach over DEA is not present in small to medium sized samples.

----
insert figure 1 about here

We see that the manufacturing productivity distribution is single-peaked throughout the period 1958-96 and with an overall decreasing dispersion and the mode moving further in the region of lower levels of productivity. The probability mass in the right tail at relatively high productivity levels beyond 0.5 is rougly constant over time.[3] The departures from the shape of a normal distribution are obvious and are confirmed by qq-plots (results not shown). It has to be noted that the lognormal model is primarily intended to illustrate the general Markov model and not as a detailed account of all peculiarities of the manufacturing productivity distribution. Furthermore, the manufacturing productivity distribution is an unconditional object and therefore subject to Brock's (1999) criticism that it is totally silent about the dynamics of the stochastic process from which it is generated.[4]

Therefore, much more interesting than the one-dimensional density estimates which summarize the information about the changes in the shape of the manufacturing productivity distribution, but cannot say anything about intra-distributional changes, would be an estimate of the transition density $\tilde{p}(\ln a_t, \ln a_{t+1})$. This object summarizes the whole information about intra-distribution dynamics and allows to read off information on (a) the changing external shape of the density, (b) the extent of intra-distribution dynamics (especially the amount of mobility and tendencies towards polarization), (c) the long-run behavior of the distribution and (d) the speed of convergence to the long-run (see Quah (1996, p. 108)).

Quah (1996, p. 117) explains how to estimate the transition density nonparametrically. The first step is a bivariate kernel density estimation of the joint density function of t and $t + k$ at the points x and y (see Wand and Jones (1995, ch. 4))

---

[3] Since the observations that do not determine the frontier function have a productivity score smaller than unity it is not surprising that most of the probability mass below the density estimates for the log productivity scores is concentrated in the range smaller than zero.

[4] In the words of Brock (1999, p. 415): "This kind of work that estimates conditional objects such as one-step-ahead conditional densities tells economists much more about the data generating process than estimation of unconditional objects such as log-normal distributions and Pareto tails."

$$\hat{f}(x,y) = \frac{1}{N(T-k)h_x h_y} \sum_{i=1}^{N} \sum_{t=1}^{T-k} K\left(\frac{x - \ln a_{it}}{h_x}\right) \cdot K\left(\frac{y - \ln a_{i,t+k}}{h_y}\right).$$

The bandwidth parameters h_x and h_y are determined by the procedure suggested in Silverman (1986, sect. 4.3.2) and the bivariate kernel density is calculated using the S-Plus routine `kde2d` from the library of Venables and Ripley (1999). Note that although the theoretical model is stated for one-period transitions primarily for convenience, it applies likewise to more general *k*-period transitions.

Integrating this joint density function in a second step gives an estimate of the period *t* marginal density $\hat{f}(x)$ and the division of the joint density by the marginal density gives the density of the log productivity levels in period $t + k$ conditional on the log productivity level in period *t*

$$\hat{p}(x,y) = \hat{f}(y \mid x) = \hat{f}(x,y)/\hat{f}(x),$$

where *x* pertains to the log productivity level in period *t* and *y* to the log productivity level in period $t + k$. Provided that the marginal density is everywhere bounded away from zero this procedure leads to a consistent estimate of the transition density function.

----
insert figure 2 about here

Figure 2 shows the nonparametric estimate of the transition density for the case $k = 5$. For a fixed productivity level on the axis representing period *t* the curve along the axis representing period $t + k$ is the conditional density of the productivity level in period $t + k$ given this particular productivity level in period *t*. The transition density is best imagined as a continuous version of a Markov transition matrix where the height of the density corresponds to the magnitude of the entries in the transition matrix (Quah (1996, p. 111)).

The dominant characteristic of the transition density in figure 2 is the marked ridge along the diagonal. Industries that have a low productivity at a certain point in time tend to have also a low productivity level five years later. A analogous statement applies to industries with high productivity levels with the exception that the transition density at high productivity levels seems to be slightly more dispersed than at low productivity levels. Thus, the clustering of most of the probability mass around the diagonal points to a substantial amount of persistence and immobility.[5]

This is confirmed by the contour plot of the transition density in figure 3 which shows slices through the transition density at the levels 0.5, 1.0, 1.5, 2.0 and 2.5. In addition to the confirmation of immobility the contour plot reveals weak evidence of polarization. At the log productivity level of about 0.5 the transition density is somewhat lower than at higher and lower productivity levels. This may be taken as a mirror image on the industry level to the finding of Quah (1996, 1997) about the evolution of the world income distribution towards bimodality.

----
insert figure 3 about here

Admittedly, in the case of manufacturing productivity distribution this finding is by far less clear cut than in the case of the world income distribution and the unconditional densities in figure 1 also do not show any tendency towards the development of a second mode. It needs a good deal of imagination to belief in twin-peaks dynamics in this case and so we do not pursue this issue further here. In addition, the finding of a unimodal transition density is quite favorable for the lognormal model which predicts a normal distribution for the log productivity levels. The finding of a large amount of persistence suggests a large value of γ_2, in the extreme case even $\gamma_2 = 1$, where productivity shocks have a permanent effect on the productivity level. In the next section we apply the recently developed unit root and stationarity tests for panel data to test the hypothesis $H_0 : \gamma_2 = 1$ more formally. By that we may provide further substantiation of the validity of our conclusion regarding the persistence of productivity in this section.

---

[5] The dominating ridge along the diagonal appears also in transition density plots for other values of *k*. The natural difference to figure 2 is that for *k* lower than 5 the plot is less dispersed, whereas it is more dispersed for values of *k* larger than 5.

6 Panel Unit Root Tests

The nonparametric estimate of the transition density in the last section points to a substantial amount of persistence in the process that translates the current into the future productivity level. This persistence is closely related to the properties of the stochastic process that governs the evolution of productivity

$$\ln a_{i,t+1} = \ln(\varphi\psi\gamma_1) + \gamma_2 \ln a_{it} + \gamma_3 \ln \bar{a}_t + u_{it}$$

which is stated here for industry i with an error term u_{it} added. The dynamic properties of this process depend to a large extent on the value of γ_2. If $\gamma_2 \in (-1, 1)$ the process is weakly stationary and will revert to its mean so that the impact of random shocks will only be transitory. On the other hand, in the unit root case $\gamma_2 = 1$ the process is a random walk (with drift) that has fundamentally different properties. Such a process is nonstationary and will never revert to its mean. Random shocks have a permanent effect on the level of the process in this case which implies that industries which have a relatively high productivity level currently will be relatively productive in future times with a high probability (see Hamilton (1994, ch. 15) for a more detailed comparison of the properties of stationary and unit root processes).

Tests for such unit roots in single time series have a long tradition in macroeconomics since the seminal work of Nelson and Plosser (1982). These tests are constructed to detect whether a stochastic process has a random walk component which is build up through the accumulation of persistent random shocks. Recently, panel versions of unit root tests that pool the time series and cross-section information in a panel have been developed to remedy for size distortions and low power that usually affect unit root tests in finite samples (see Banerjee (1999) for a survey). The purpose of these panel tests is to give a summary of the prevalence of unit root or stationary time series in a panel of data. Below we describe one standard test for the null hypothesis of a unit root panel proposed by Im et al. (2002) and one test for the null of a stationary panel proposed by Hadri (2000).

The panel unit root test of Im et al. (2002), henceforth IPS test, is intended to test the null hypothesis of a unit root for a sample of N cross-section units that are observed over T time periods. The test is based on the estimation of the familiar ADF regression (Dickey and Fuller 1979)

$$\Delta z_{it} = \alpha_i + (\delta_i t) + \rho_i z_{i,t-1} + \Sigma_{j=1}^{p_i} \theta_{ij} \Delta z_{i,t-1} + u_{it} \; ; t = 1, ..., T$$

for each cross-section unit $i = 1, ..., N$ separately. Here Δ denotes the first difference operator and z_{it} is the result of the transformation $z_{it} = \ln a_{it} - \frac{1}{N}\Sigma_{i=1}^{N} \ln a_{it}$ to eliminate time effects. The deterministic part of the regression is represented by an intercept and a time trend where the latter needs not to be included in the regression. How many lagged differences of z_{it} are included in order to account for autocorrelation is determined for each cross-section separately by testing down from a maximum lag length of 5 periods until the coefficient pertaining to the highest lag is rejected on a 5% level of significance. To test the unit root hypothesis $H_0 : \rho_i = 0$ for all i against $H_1 : \rho_i < 0$ for at least one i, the t-statistic of ρ_i is calculated for each individual cross-section unit i. The IPS test is then constructed by taking the average of all individual ADF t-statistics $\bar{t}$ which is shown to follow the probability law $\sqrt{N}\,(\bar{t} - \mu_{\bar{t}})/\sigma_{\bar{t}} \xrightarrow{d} N(0, 1)$ as$N \to \infty$ and $T \to \infty$, where $\xrightarrow{d}$ denotes convergence in distribution. $\mu_{\bar{t}}$ and $\sigma_{\bar{t}}$ are the averages of the means and variances of the individual ADF t-statistics under the null hypothesis which are computed by simulation and tabulated in Im et al. (2002, table 3) for different values of *T*, lag lengths p_i and different specifications of the deterministic part.[6]

This procedure permits a large amount of heterogeneity between the cross-section units by the removal of time effects, different lag lengths for the lagged differences and different coefficients in each ADF regression. The removal of the time effects has an additional fortunate consequence in this application because it also eliminates the effect of the cross-section invariant spillover term $\ln \bar{a}_t$ from the law of motion of $\ln a_t$. This term is essentially part of the drift component in the random walk process for $\ln a_t$.

The reverse null hypothesis, that of stationarity, can be tested by the test of Hadri (2000) which builds on results of Nabeya and Tanaka (1988) and Kwiatkowski et al. (1992). The basic model is here $z_{it} = \alpha_i + \delta_i t + r_{it} + u_{it}$ for $z_{it} = \ln a_{it} - \frac{1}{N}\Sigma_{i=1}^{N} \ln a_{it}$, where $r_{it} = r_{i,t-1} + \varepsilon_{it}$ and

[6] This is the alternative standardized t-bar statistic, termed W_{tbar} in Im et al. (2002, p. 10).

the errors $\varepsilon_{it} \sim iid(0, \sigma_\varepsilon^2)$ (again the inclusion of an intercept is adviced but the deterministic trend may be neglegted). Inserting the unit root assumption for r_{it} we arrive at

$$z_{it} = \alpha_i + (\delta_i t) + \Sigma_{\tau=1}^{t} \varepsilon_{i\tau} + u_{it} = \alpha_i + (\delta_i t) + e_{it},$$

where e_{it} contains the stochastic components $\Sigma_{\tau=1}^{t} \varepsilon_{i\tau} + u_{it}$. The stationary null hypothesis can be stated as $H_0 : \sigma_\varepsilon^2/\sigma_u^2 = 0$ whereas the alternative is $H_1 : \sigma_\varepsilon^2/\sigma_u^2 > 0$. Therefore the test statistic is computed by

$$LM = \frac{1}{N} \sum_{i=1}^{N} \frac{\frac{1}{T^2} \Sigma_{t=1}^{T} S_{it}^2}{\hat{\sigma}_{\varepsilon,i}^2},$$

where $S_{it} = \Sigma_{\tau=1}^{t} \hat{u}_{i\tau}$ is the partial sum procress of the residuals of a time series regression of y_{it} on a constant (and a time trend if appropriate) $\hat{u}_{it} = y_{it} - \hat{\alpha}_i - (\hat{\delta}_i t)$ and the long-run variance $\hat{\sigma}_\varepsilon^2$ is calculated by the Newey-West estimator (Newey and West (1987))

$$\hat{\sigma}_{\varepsilon,i}^2 = \frac{1}{N} \sum_{i=1}^{N} \left[\hat{\gamma}_{0i} + 2 \sum_{j=1}^{p} \left(1 - \frac{j}{p+1}\right) \hat{\gamma}_{ji} \right], \text{ where } \hat{\gamma}_{ji} = \frac{1}{T} \sum_{\tau=j+1}^{T} \hat{u}_{it} \hat{u}_{i,t-j}.$$

The lag length is here chosen by the Newey-West (1994) proposal $p = \lfloor 4(T/100)^{2/9} \rfloor$ where $\lfloor \cdot \rfloor$ denotes the integer part. The statistic LM is then transformed to an asymptotic standard normal random variate $\sqrt{N}(LM - \mu_{LM})/\sigma_{LM} \xrightarrow{d} N(0,1)$ as $N \to \infty$. The mean and standard deviation of LM, μ_{LM} and σ_{LM}, have been derived analytically by Hadri and Larsson (2002) in dependence of a finite time dimension T for the specification with an intercept only and the specification with both intercept and trend.

The simultaneous application of an unit root test and a test for stationarity allows for confirmatory analysis. By this procedure we obtain more reliable conclusions concerning the unit root nature to productivity if the IPS test is not able to reject its unit root null while the Hadri test rejects its stationary null. Choi (2000) investigates the properties of confirmatory analysis in panel data by a Monte Carlo exercise and finds that besides the increased reliability of inference in the confirmatory cases, confirmatory analysis has also the potential to detect cases in which the panel consists of stationary as well as nonstationary time series. Such cases of mixed structure are indicated by a simultaneous rejection of the respective null hypotheses by both the unit root and the stationarity test.

The resulting unit root test statistics and their p-values (in parentheses) for the manufacturing industries on the three- and four-digit level of aggregation together with various subsamples of the three-digit sample are summarized in table 1. It is important for the interpretation of the statistics and p-values that the IPS tests are lower-tailed tests whereas the Hadri tests are upper-tailed tests.

Table 1
Unit Root and Stationarity Tests for the Productivity Scores

	IPS (intercept)	IPS (trend)	Hadri (intercept)	Hadri (trend)
Three-Digit Level	5.2515 (1.0000)	2.4709 (0.9933)	38.2100 (0.0000)	24.2225 (0.0000)
Four-Digit Level	7.3350 (1.0000)	-2.9154 (0.0018)	72.6056 (0.0000)	40.0485 (0.0000)
Period 1958-73	1.1421 (0.8733)	-10.9480 (0.0000)	21.4127 (0.0000)	13.6635 (0.0000)
Period 1974-96	3.9549 (1.0000)	0.2099 (0.5831)	28.3633 (0.0000)	18.1229 (0.0000)
High-Tech Industries	5.2130 (1.0000)	4.0280 (1.0000)	18.8165 (0.0000)	15.7811 (0.0000)
Low-Tech Industries	2.6222 (0.9956)	-0.1395 (0.4445)	33.3421 (0.0000)	18.1977 (0.0000)
Durable Goods	5.3092 (1.0000)	2.1161 (0.9828)	28.5296 (0.0000)	16.6648 (0.0000)
Nondurable Goods	2.0007 (0.9773)	1.5350 (0.9376)	23.0917 (0.0000)	18.6401 (0.0000)

Note: all tests are based on logged data with fixed time effects removed; p-values are in parentheses.

The results show that in general the IPS tests fail to reject their unit root hypotheses whereas all Hadri tests strongly reject their stationary null hypotheses. This holds for the specification of the deterministic part as only an intercept and with two exceptions for the specification with both intercept and trend. The two exceptions are the IPS tests with trend for the four-digit industries and for the period 1958-73 prior to the onset of the productivity slowdown. The pattern of results is also robust to splitting the sample into subsamples of high-tech and low-tech industries according to Hadlock et al. (1991)[7] and industries that produce durable and nondurable goods according to Quah and Sargent (1993)[8], respectively.

[7] Hadlock et al. (1991) classigy the three-digit manufacturing industries as high-tech if the industry's proportion of R&D employment in the year 1989 was at least equal to the average proportion of all industries surveyed and as low-tech otherwise. The result of this procedure are 34 high-tech and 106 low-tech industries. This classification has also been used in the recent empirical literature on the industry life cycle (see e.g. Agarwal and Audretsch (2001) and Agarwal and Gort (1996)).

[8] In appendix B of Quah and Sargent (1993) a classification of two-digit industries in those that produce durable

Thus, we obtain a clear confirmation of a unit root in the stochastic process that governs the dynamics of the total factor productivity of the U.S. manufacturing industries. This supports the visual impression that we have already gained from the nonparametric estimate of the transition kernel and lends support to the estimate $\hat{\gamma}_2 = 1$ in the lognormal specification. Bernard and Jones (1996) also use panel unit root tests to assess the convergence of different sectors across 14 OECD countries. They interpret the nonrejection of a unit root in the difference of total factor productivity of the manufacturing sector in a country above the median across all countries as evidence for nonconvergence in that sector. In services the unit root could be rejected leading to the conclusion of convergence in the service sector.

The finding of a unit root in the process of $\ln a_{it}$ implies also a unit root in the processes for mean μ_t and variance σ_t^2. Results for the efficient unit root test DF-GLS of Elliott et al. (1996) and the stationarity test KPSS of Kwiatkowski et al. (1992) applied to the time series of the cross-section means and variances of the log productivity scores of the three-digit manufacturing industries are shown in table 2.[9] The lag length is again chosen by the Newey-West (1994) proposal and results in a value of 3 for the 39 time series observations.

Table 2
DF-GLS and KPSS Tests for Mean and Variance of the Log Productivity Score

	DF-GLS (intercept)	DF-GLS (trend)	KPSS (intercept)	KPSS (trend)
mean	-1.9343*	-1.7933	0.6378**	0.1462*
variance	-2.2308**	-2.0123	0.8033***	0.1367*

Note: rejections on 1%, 5% and 10% level are marked by ***, ** and *, respectively; critical values are obtained from Hamilton (1994, table B.6, case 1) for DF-GLS (intercept), Elliott et al. (1996, table I) for DF-GLS (trend) and Sephton (1995) for both KPSS tests.

For both mean and variance we observe (weak) rejections of the unit root null in the intercept case but no rejections in the trend case. The stationarity null hypothesis is in any case rejected, although only on 10 percent level in the case with trend when we rely on the finite-sample critical values based on Sephton's (1995) response surface estimates. Here, the small sample size in the time series case together with the requirement to estimate an additional parameter

and those that produce nondurable goods is reported. We assign all three-digit industries to durable good or nondurable goods according to the classification of the two-digit industry to which they pertain. This results in 72 and 68 three-digit industries that produce durable and nondurable goods, respectively.

9 See Maddala and Kim (1998, pp. 113f. and pp. 120ff.) for good textbook expositions of both tests.

may cause a lack of power. On the whole, this pattern of results also implies that γ_2 exhibits a value that is at least quite close to unity and is therefore consistent with the findings of the panel unit root and stationarity tests.

7 Conclusion

The analyses reported in this paper investigate the dynamics of the U.S. manufacturing productivity distribution both theoretically and empirically. The theoretical model is based on a Markov process where the technological level of an industry determines its investment in research activities that together with inter-industry spillover effects stochastically affect next periods productivity. The strength of these effects is causal for the amount of persistence with which todays productivity translates into tomorrows productivity. Using a mixture of nonparametric and parametric econometric techniques we have found that the most important dynamic feature of the productivity scores is persistence. This finding implies that once gained productivity levels are unlikely to be quickly eroded, but it also implies that it is difficult to catch up from low to high productivity levels.

Persistence becomes clearly visible in the plot of the transition density which is characterized by a dominating ridge along the diagonal. For the interpretation of all above reported unit root test results it has to be emphasized that the unit root finding need not necessarily imply that γ_2 is exactly equal to one but just that it has a value very close to one. According to the insights provided by Blough (1992) it is logically impossible to distinguish a stationary, but highly persistent, process from a random walk in finite samples. Against this background we interpret the above results as indicating that persistence plays a major role in the productivity developement of the U.S. manufacturing sector. This persistence of productivity is equivalent to the statement that there is a high probability of staying in a neighborhood of the current productivity level in the future relative to the probability of reaching a productivity level that is considerably smaller or larger than the current one. In that sense success-breeds-success dynamics have force not only on the firm level as many studies have found, but also on the level of three- or four-digit industries.

The empirical findings are altogether consistent with the dynamic behavior of the lognormal specification of the general Markov model. The conditional densities that compose the transition kernel look roughly normal, although the shapes of the unconditional densities are very different from the shape of a normal density. In this respect the lognormal specification fails, but it should be taken into consideration that this specification is chosen mainly for illustrative purposes and because the transition density and the law of motion of the density functions can be derived analytically. Future research will focus on other elaborations of the general Markov model to adress the dynamics of higher moments of the manufacturing productivity distribution. Those specifications will probably not be manageable analytically but instead will require the application of simulation methods.

References

Agarwal, R., Audretsch, D.B. (2001), Does Entry Size Matter? The Impact of the Life Cycle and Technology on Firm Survival, Journal of Industrial Economics, vol. 49, pp. 21-43.

Agarwal, R., Gort, M. (1996), The Evolution of Markets and Entry, Exit and Survival of Firms, Review of Economics and Statistics, vol.78, pp. 489-498.

Andersen, P., Petersen, N.C. (1993), A Procedure for Ranking Efficient Units in Data Envelopment Analysis, Management Science, vol. 39, pp. 1261-1264.

Arthur, B.W. (1989), Competing Technologies, Increasing Returns, and Lock-in by Historical Events, Economic Journal, vol. 99, pp. 116-131.

Banerjee, A. (1999), Panel Data Unit Roots and Cointegration: An Overview, Oxford Bulletin of Economics and Statistics, vol. 61, pp. 607-629.

Banker, R.D., Gadh, V.M., Gorr, W.L. (1993), A Monte Carlo Comparison of Two Production Frontier Estimation Methods: Corrected Ordinary Least Squares and Data Envelopment Analysis, European Journal of Operational Research, vol. 67, pp. 332-343.

Bartelsman, E.J., Gray, W. (1996), The NBER Manufacturing Productivity Data Base, NBER Technical Working Paper 205.

Bernard, A.B., Jones, C.I. (1996), Productivity Across Industries and Countries: Time Series Theory and Evidence, Review of Economics and Statistics, vol. 78, pp. 135-146.

Billingsley, P. (1995), Probability and Measure, 3rd ed., New York: Wiley.

Blough, S.R. (1992), The Relationship between Power and Level for Generic Unit Root Tests in Finite Samples, Journal of Applied Econometrics, vol. 7, pp. 295-308.

Brock, W.A. (1999), Scaling in Economics: A Reader's Guide, Industrial and Corporate Change, vol. 8, pp. 409-446.

Caves, R.E. (1998), Industrial Organization and New Findings on the Turnover and Mobility of Firms, Journal of Economic Literature, vol. 36, pp. 1947-1982.

Charnes, A., Cooper, W.W., Lewin, A.Y., Seiford, L.M. (eds.) (1994), Data Envelopment Analysis: Theory, Methodology and Applications, Boston: Kluwer.

Choi, C.-Y. (2000), Confirmatory Analysis in Panel Unit Root Tests, mimeo.

Dickey, D.A., Fuller, W.A. (1979), Distribution of the Estimators for Autoregressive Time Series with a Unit Root, Journal of the American Statistical Association, vol. 74, pp. 427-431.

Durlauf, S.N. (1993), Nonergodic Economic Growth, Review of Economic Studies, vol. 60, pp. 349-366.

Durlauf, S.N., Quah, D.T. (1999), The New Empirics of Economic Growth, in: J.B. Taylor, M. Woodford (eds.), Handbook of Macroeconomics, vol. 1A, Amsterdam: Elsevier, pp. 235-308.

Elliott, G., Rothenberg, T.J., Stock, J.H. (1996), Efficient Tests for an Autoregressive Unit Root, Econometrica, vol. 64, pp. 813-836.

Ericson, R., Pakes, A. (1995), Markov-Perfect Industry Dynamics: A Framework for Empirical Work, Review of Economic Studies, vol. 62, pp. 53-82.

Feller, W. (1971), An Introduction to Probability Theory and Its Applications, vol. 2, 2nd ed., New York: Wiley.

Gallant, A.R. (1997), An Introduction to Econometric Theory, Princeton: Princeton University Press.

Greene, W.H. (1993), The Econometric Approach to Efficiency Analysis, in: H.O. Fried, C.A.K. Lovell, S.S. Schmidt (eds.), The Measurement of Productive Efficiency, Oxford: Oxford University Press, pp. 68-119.

Hadlock, P., Hecker, D., Gannon, J. (1991), High Technology Employment: Another View, Monthly Labor Review, vol. 114, pp. 26-30.

Hadri, K. (2000), Testing for Stationarity in Heterogeneous Panel Data, Econometrics Journal, vol. 3, pp. 148-161.

Hadri, K., Larsson, R. (2002), Testing for Stationarity in Heterogeneous Panel Data Where the Time Dimension is Finite, Working Paper, University of Liverpool.

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

Im, K.S., Pesaran, M.H., Shin, Y. (2002), Testing for Unit Roots in Heterogeneous Panels, Journal of Econometrics, forthcoming.

Johnson, N.L., Kotz, S. (1970), Distributions in Statistics: Continuous Univariate Distributions-1, New York: Wiley.

Jones, C.I. (1995), R&D-Based Models of Economic Growth, Journal of Political Economy, vol. 103, pp. 759-784.

Jones, M.C., Marron, J.S., Sheather, S.J. (1996), A Brief Survey of Bandwidth Selection for Density Estimation, Journal of the American Statistical Association, vol. 91, pp. 401-407.

Jorgenson, D.W. (1990), Productivity and Economic Growth, in: E.R. Berndt, J.E. Triplett (eds.), Fifty Years of Economic Measurement, Chicago: University of Chicago Press, pp. 19-118.

Kwiatkowski, D., Phillips, P.C.B., Schmidt, P., Shin, Y. (1992), Testing the Null Hypothesis of Stationarity Against the Alternative of a Unit Root, Journal of Econometrics, vol. 54, pp. 159-178.

Lehmann, E.L., Casella, G. (1998), Theory of Point Estimation, 2nd ed., New York: Springer.

Loury, G.C. (1981), Intergenerational Transfers and the Distribution of Earnings, Econometrica, vol. 49, pp. 843-867.

Maddala, G.S., Kim, I.-M. (1998), Unit Roots, Cointegration, and Structural Change, Cambrigde (Mass.): Cambridge University Press.

Maskin, E., Tirole, J. (1988), A Theory of Dynamic Oligopoly, I: Overview and Quantity Competition with Large Fixed Costs, Econometrica, vol. 56, pp. 549-569.

Maskin, E., Tirole, J. (2001), Markov Perfect Equilibrium: I. Observable Actions, Journal of Economic Theory, vol. 100, pp. 191-218.

Morrison, C.J., Siegel, D. (1997), External Capital Factors and Increasing Returns in U.S. Manufacturing, Review of Economics and Statistics, vol. 79, pp. 647-654.

Nabeya, S., Tanaka, K. (1988), Asymptotic Theory of a Test for the Constancy of Regression Coefficients Against the Random Walk Alternative, Annals of Statistics, vol. 16, pp. 218-235.

Nelson, C.R., Plosser, C.I. (1982), Trends and Random Walks in Macroeconomic Time Series, Journal of Monetary Economics, vol. 10, pp. 139-162.

Newey, W.K., West, K.D. (1987), A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix, Econometrica, vol. 55, pp. 703-708.

Newey, W.K., West, K.D. (1994), Automatic Lag Selection for Covariance Matrix Estimation, Review of Economic Studies, vol. 61, pp. 631-653.

Quah, D. (1996), Convergence Empirics Across Economies with (Some) Capital Mobility, Journal of Economic Growth, vol. 1, pp. 95-124.

Quah, D. (1997), Empirics for Growth and Distribution: Stratification, Polarization, and Convergence Clubs, Journal of Economic Growth, vol. 2, pp. 27-59.

Quah, D., Sargent, T.J. (1993), A Dynamic Index Model for Large Cross Sections, in: J.H. Stock, M.W. Watson (eds.), Business Cycles, Indicators, and Forecasting, University of Chicago Press, NBER, pp. 285-309.

Romer, P.M. (1986), Increasing Returns and Long-Run Growth, Journal of Political Economy, vol. 94, pp. 1002-1037.

Ruggiero, J. (1999), Efficiency Estimation and Error Decomposition in the Stochastic Frontier Model: A Monte Carlo Analysis, European Journal of Operational Research, vol. 115, pp. 555-563.

Septhon, P.S. (1995), Response Surface Estimates for the KPSS Stationarity Test, Economics Letters, vol. 47, pp. 255-261.

Sheather, S.J., Jones, M.C. (1991), A Reliable Data-Based Bandwidth Selection Method for Kernel Density Estimation, Journal of the Royal Statistical Society, Series B, vol. 53, pp. 683-690.

Silverman, B.W. (1986), Density Estimation for Statistics and Data Analysis, London: Chapman&Hall, 1986.

Stokey, N.L., Lucas, R.E. (1989), Recursive Methods in Economic Dynamics, Cambridge (Mass.): Harvard University Press.

Sutton, J. (1997), Gibrat's Legacy, Journal of Economic Literature, vol. 35, pp. 40-59.

Venables, W.N., Ripley, B.D. (1999), Modern Applied Statistics with S-Plus, 3rd ed., New York: Springer.

Wand, M.P., Jones, M.C. (1995), Kernel Smoothing, London: Chapman&Hall.

Figure 1
Univariate Kernel Density Plots for the Productivity Scores

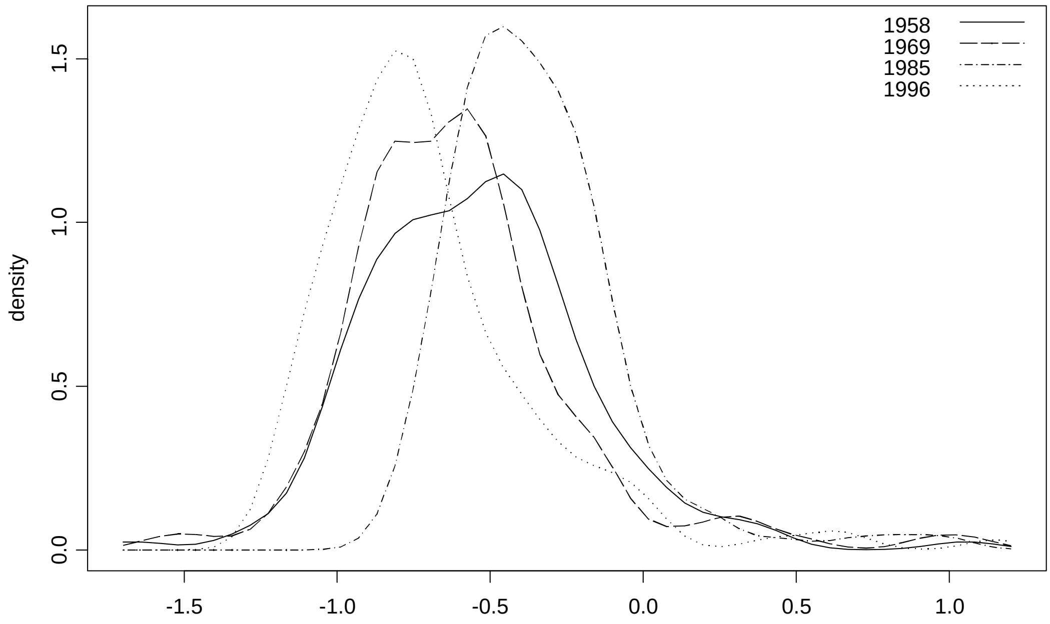


Figure 2
Nonparametric Estimate of the Markov Transition Kernel

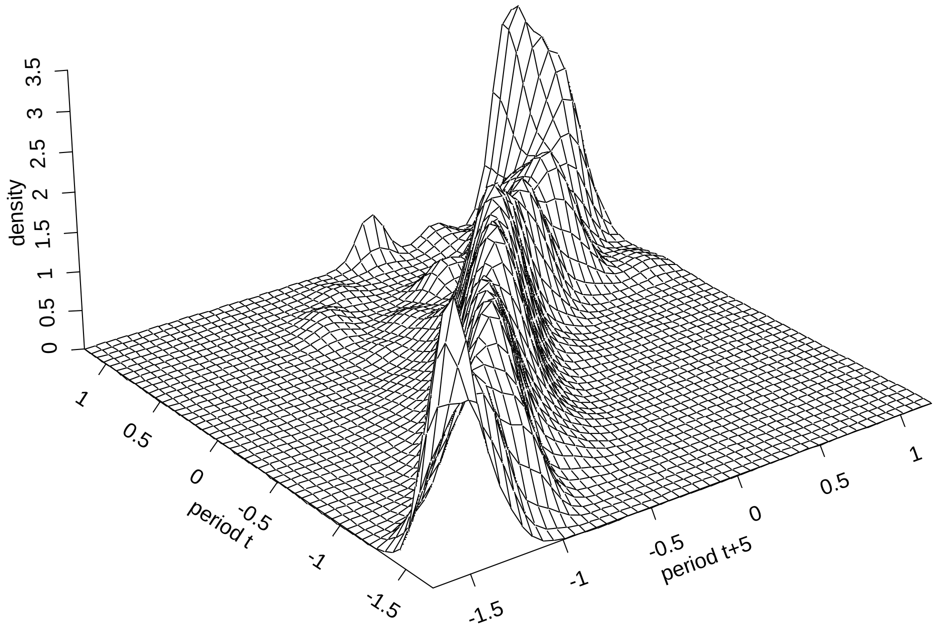
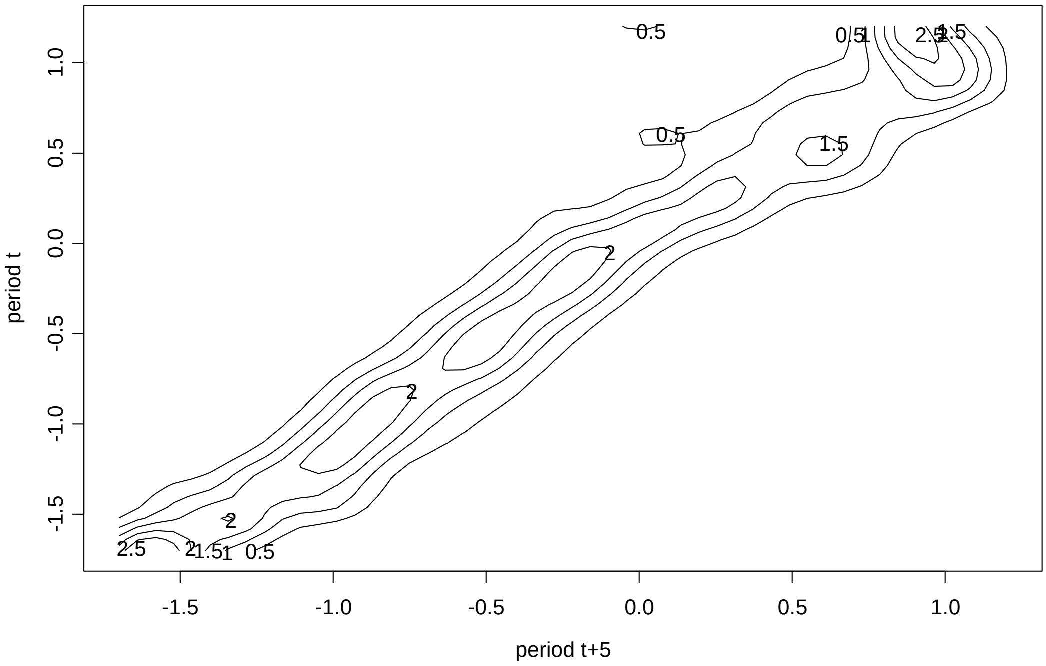


Figure 3
Contour Plot of the Markov Transition Kernel of Figure 2



Jenaer Schriften zur Wirtschaftswissenschaft

2003

1/2003 Wolfgang Kürsten: Synergetische Merger, Co-Insurance und Shareholder Value, oder: Wer profitiert von "wertschaffenden" Fusionen?

2/2003 Roland Helm, Laura Manthey, Armin Scholl und Michael Steiner: Empirical Evaluation of Preference Elicitation Techniques from Marketing and Decision Analysis.

3/2003 Wolfgang Kürsten: Grenzen und Reformbedarfe der Sicherheitsäquivalentmethode in der (traditionellen) Unternehmensbewertung. Erwiderung auf die Anmerkungen von Ralf Diedrich und Jörg Wiese in der ZfbF.

4/2003 Uwe Cantner, Holger Graf: Cooperation and Specialization in German Technology Regions.

5/2003 Jens J. Krüger: On the Dynamics of the U.S. Manufacturing Productivity Distribution.

6/2003 Uwe Cantner, Dirk Fornahl und Holger Graf: Innovationssystem und Gründungsgeschehen in Jena. Erste Erkenntnisse einer Unternehmensbefragung.

7/2003 Peter Kischka: Faktorzuweisungen im Rubin-Modell.

2002

Die Schriftenreihe war bis 2002 in Reihe A "Betriebswirtschaftslehre" und Reihe B "Volkswirtschafts-lehre" unterteilt.

Reihe A:

1/2002 Dorothea Alewell und Petra Koller: Die Sicherung von Humankapitalinvestitionen über Rückzahlungsklauseln – Ergebnisse einer empirischen Untersuchung. *Erschienen in: Mitteilungen aus der Arbeitsmarkt- und Berufsforschung 1/2002, S. 107-122.*

2/2002 Dorothea Alewell und Petra Koller: Der Einsatz und die Bewertung von Rückzahlungsklauseln für Weiterbildungsmaßnahmen durch Unternehmen - Ergebnisse einer empirischen Untersuchung. *Erscheint in: Zeitschrift für Personalforschung 2003*

3/2002 Martin Kloyer: Begrenzung des Opportunismus in FuE-Lieferbeziehungen - Eine empirische Untersuchung.

4/2002 Martin Kloyer: Patentierungsstrategien: Bestimmung des Anspruchsumfangs.

5/2002 Armin Scholl und Lutz Häselbarth: Zielgerichtete Koordination mit Hilfe der Dantzig-Wolfe-Dekomposition - Eine betriebswirtschaftliche Fallstudie. *Erscheint in: WISU - das Wirtschaftsstudium 32 (2003).*

Reihe B:

1/2002 Markus Pasche: Heterogeneous Behavioral Rules in the Oligopolistic Case.

2/2002 Markus Pasche: Playing Fair: Rationality and Norm-guided Behavior in Games.

3/2002 Uwe Cantner und Jens J. Krüger: Geroski's Stylized Facts and Mobility in Large German Manufacturing Firms.
